Scenario, analysis, and design of privacy throughout life demonstrator

C Roosendaal, A.P.; Borcea-Pfitzmann, K.; Steinbrecher, S.; Storf, K.; Hansen, M.; Raguse, M.; Kuczerawy, A.; Wouters, K.; Pfitzmann, A.; Böhme, R.; Berthold, S.; Dobias, J.

Publication date:
2011

Document Version
Peer reviewed version

Scenario, Analysis, and Design of Privacy Throughout Life Demonstrator

Editors:	Katrin Borcea-Pfitzmann (TUD)
Reviewers:	Hans Hedbom (KAU)
	Peter Wolkerstorfer (CURE)

Identifier:	D1.3.1
Type:	Deliverable
Version:	1.0
Class:	Public
Date:	February 28, 2011

Abstract

The main contribution of this deliverable to the research field of Privacy-Enhancing Identity Management Throughout Life consists in a comprehensive analysis of requirements. Those requirements comprise high-level requirements regarding issues of transparency, data minimisation, controlled data processing, user-controlled identity management, delegation, practicability, and change management. Further, more specific requirements from the socio-cultural and delegation points of view as well as from the actual nature of the envisaged demonstrator (which is backup and synchonisation) are being elaborated.

Apart from the elaboration of requirements, solutions based on specific tools and mechanisms are described and discussed. This includes a list of recommendations for policy makers specially addressing lifetime aspects of privacy and identity management. In addition, this documents provides a extensive glossary of terms and concepts important to the given research field.

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 216483 for the project PrimeLife.

Members of the PrimeLife Consortium

1.	IBM Research GmbH	IBM	Switzerland
2.	Unabhängiges Landeszentrum für Datenschutz	ULD	Germany
3.	Technische Universität Dresden	TUD	Germany
4.	Karlstads Universitet	KAU	Sweden
5.	Università degli Studi di Milano	UNIMI	Italy
6.	Johann Wolfgang Goethe - Universität Frankfurt am Main	GUF	Germany
7.	Stichting Katholieke Universiteit Brabant	TILT	Netherlands
8.	GEIE ERCIM	W3C	France
9.	Katholieke Universiteit Leuven	K.U.Leuven	Belgium
10.	Università degli Studi di Bergamo	UNIBG	Italy
11.	Giesecke & Devrient GmbH	GD	Germany
12.	Center for Usability Research & Engineering	CURE	Austria
13.	Europäisches Microsoft Innovations Center GmbH	EMIC	Germany
14.	SAP AG	SAP	Germany
15.	Brown University	UBR	USA

List of Contributors

Contributions from several PrimeLife partners are contained in this document. The following list presents the contributors for the chapters of this deliverable.

Chapter	**Author(s)**
Chapter 1: Introduction	Arnold Roosendaal (TILT), Katrin Borcea-Pfitzmann (TUD)
Chapter 2: Privacy and Identity Management throughout Life	Arnold Rosendaal (TILT), Sandra Steinbrecher (TUD), Andreas Pfitzmann (TUD), Katrin Borcea-Pfitzmann (TUD), Katalin Storf (ULD), Marit Hansen (ULD)
Chapter 3: Requirements and Concepts for Privacy-Enhancing Daily Life	Katalin Storf (ULD), Marit Hansen (ULD), Maren Raguse (ULD), Arnold Roosendaal (TILT), Aleksandra Kuczerawy (K.U.Leuven), Karel Wounters (K.U.Leuven) (Sec. 3.1 High-Level requirements for Privacy Throughout Life), Sandra Steinbrecher (TUD), Andreas Pfitzmann (TUD), Rainer Böhme (TUD), Stefan Berthold (TUD), Marit Hansen (ULD) (Sec. 3.2 Tools and Mechanisms)
Chapter 4: Demonstrator to Show the Interplay Between Scenarios	Jaromir Dobiáš (TUD), Katrin Borcea-Pfitzmann (TUD)
Chapter 5: Conclusion	Katrin Borcea-Pfitzmann (TUD) (Sec.5.1 Lessons Learned), Marit Hansen (ULD), Katalin Storf (ULD) (Sec. 5.2 Recommendations for Policy Makers)

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Privacy and Identity Management has been discussed from very different points of view in the past (cf. [CK01, vdBL10, CSS+05, LSH08] etc.) and it is still subject to research in specific research projects such as PrimeLife[1], PICOS[2], GINI-SA etc. One important aspect in researching privacy and identity management had been neglected so far, however, namely – the consideration of the peculiarities of a human being's life and his perceptions and abilities regarding his privacy management. This is what this deliverable deals with. It frames the whole research area by studying stages of life, dynamics of life, and areas of life as well as their relationships to the privacy and identity management by an individual. Also it discusses requirements to be considered when developing solutions that tackle the challenges of lifelong privacy and identity management. Finally, this document gives an overview of a selected demonstrator that takes up the indicated challenge and that provides additional findings generally valid for the whole field of research.

The moment that a formal identity is created is usually at the birth of an individual. However, privacy and identity management related issues already take place before birth. During the future mother's pregnancy, files are created containing information on hereditary characteristics and the development of the foetus. Furthermore, information about the family of the unborn child is collected and insurances need to be taken out.

A similar process takes place after decease of an individual. Identity does not terminate immediately after death, but rather decays over time as rights and obligations terminate. For the purpose of pension funds and life insurances, the identity remains for a significant period. Besides, the personal details of the deceased person will remain accessible in municipal registers for historical purposes.

Having such specifics in mind, the topic will be approached first from a rather general point of view, i.e., the concept of identity is looked at from different points of view describing where (formal) identities are established and what their functions are. Following that, a comprehensive analysis is conducted aiming at determining requirements particularly valid within the setting of lifelong privacy. The requirements will be discussed by applying them to the demonstrator that will be implemented within workpackage WP1.3

[1] http://www.primelife.eu/
[2] http://www.picos-project.eu

and technically described in deliverable D1.3.2.

1.1 Privacy and Identity Management

Privacy and identity management are really broad concepts. This is why we focus on formal identities of individuals. These two concepts are closely related, but the idea is that identity management in formal contexts is a necessary condition for adequate protection of privacy[3] of individuals. Keeping contexts separated and having control over what data are disclosed to whom can be facilitated by proper identity management, when different (partial) identities can be used for different contexts. Identities can differ depending on the contexts they are used in. For instance, specific aspects of one's identity may be more relevant than others according to the purpose and use of the formal identity.

In order to give a comprehensive overview of the relevance of formal identities and the management of these identities, four specific contexts in which formal identities play a role are described, namely government, health care, education, and employment.

Even though the focus is on formal identities – described from the perspective of a number of EU countries – there is also attention for informal identities to provide the entire spectrum of privacy and identity management issues throughout life. Several domains are described and specific issues are touched upon.

1.2 Throughout Life

The four chosen key areas regarding formal identities describe identity management throughout life. It should be noted that the lifespan of a person's identity extends beyond their life. Wherever relevant, the identity establishment and use before life and after decease are therefore also described.

Furthermore, a number of questions arise when looking at identity and privacy from a lifespan perspective:

- How can a child after birth, a minor or a mentally challenged person manage their identities?
- How can a person delegate consent to such collection and processing, and still be "informed" as the law demands?
- How can they consent to collection or processing of information on their identity?
- How can we qualify the sensitivity of identity information from a balanced or fair perspective, when we are unable to ask the person(s) involved?

The problems that arise from these issues are described in this deliverable.

[3] When talking about privacy, we refer to the definition given in [BBP11]: "Privacy of a physical entity is the result of negotiating and enforcing when, how, to what extent, and in which context which of its data is disclosed to whom."

1.3 Structure of this Deliverable

The contents of this deliverable represents a summary of the results documented in the corresponding PrimeLife heartbeats.

Accordingly, Chapter 1 has input from heartbeat "H1.3.3 Analysis of privacy and identity management throughout life" [RSH+09] framing the research topic and ranging it in the overall research area of privacy and identity management.

Chapter 2 deals with general issues of the concept of *identity* as well as of terminology related to lifelong privacy management. Both of the are parts of the aforementioned PrimeLife heartbeat "H1.3.3 Analysis of privacy and identity management throughout life" [RSH+09].

Chapter 3 bases on PrimeLife heartbeat "H1.3.5 Requirements and Concepts for Identity Management throughout Life" [SHP+09] specifically elaborating on high-level requirements for lifelong privacy as well as on descriptions of tools and mechanisms enabling lifelong privacy.

In Chapter 4 specifically deals with the development of a demonstrator designated to present the main features of lifelong privacy and identity management. First, a variety of prototype ideas are presented and discussed. These were fleshed out within the frames of PrimeLife heartbeat "H1.3.4 Definition of: Prototype ideas for selected scenarios " [BRS+09]. Also, more specific requirements are defined that take the chosen demonstrator into account and reflect on the actual nature, social-cultural and delegation-related issues of the demonstrator. The latter were depicted in "H1.3.7 Second thoughts on the WP 1.3 demonstrator" [HHB+10]. Determining solutions for the most relevant requirements was the objective of PrimeLife heartbeat "H1.3.6 Towards a Privacy-Enhanced Backup and Synchronisation Demonstrator Respecting Lifetime Aspects" [DB10]. The same chapter discusses further scenarios and use cases of the privacy-enhanced backup and synchronisation demonstrator, which have been described in PrimeLife heartbeat "H1.3.7 Second thoughts on the WP 1.3 demonstrator" [HHB+10].

The document summarises the findings in Chapter 5 and takes up the definition of recommendations for policy makers given in PrimeLife heartbeat "H1.3.5 Requirements and Concepts for Identity Management throughout Life" [SHP+09].

Remark: For the purpose of readability we refrain from using gender-neutral pronouns such as "he/she". Accordingly, gendered pronouns are used in a non-discriminatory sense and are meant to represent both genders.

Chapter 2

Privacy and Identity Management throughout Life

2.1 Basics of the Concept of Identity

2.1.1 General Aspects of Identity

When talking about identity management, it is necessary to first have an idea of what identity is. This section briefly describes identity from both a social science and a technical perspective. It also discusses some concepts related to identities in the digital world. Individuals interact with other individuals and organisations in many different relations, all of which are connected to different roles of the individual. Goffman defines identity as "the result of publicly validated performances, the sum of all roles played by the individual, rather than some innate quality." [Gof59] In this respect, all different roles can be seen as (partial) identities.

Depending on the context (relation) between the individual and the person or entity they interact with, certain information is disclosed or not. The information disclosed and characteristics associated to the individual are attributes of this individual. Individuals from a data perspective can therefore be seen as a (large) collection of attributes. For a concrete partial identity the attributes take specific values. So 'first name' is an attribute label while 'Peter' is an attribute value.

"Different (kinds of) relationships involve different kinds of information constituting the individual's identity. A single individual therefore consists of different characterisations tied to the different contexts in which she operates. For example, the co-workers in a work-related context will characterise an individual differently than the friends that interact with the same individual in the context of friendship. The relevant attributes associated to an individual are different in a working environment than in a social environment and individuals may also represent themselves differently throughout such contexts." [CLS11, p.24] Some attributes may thus take different values in different context. For instance, James' nickname may be 'Jim' among his friends, whereas his colleagues might call him 'Captain Slow' (behind his back).

According to [Gof59], different contexts impose different rules on behaviour and people play different roles (as in a theatre play) in different contexts. Also they present different faces of themselves. Thus, we may say that individuals give different performances in everyday life. Audience segregation is at the same time a natural effect and an important enabler of the part one performs. "[B]y audience segregation the individual ensures that those before whom he plays one of his parts will not be the same individuals before whom he plays a different part in another setting." [Gof59] Audience segregation is a device for protecting fostered impressions. Rachels states that this audience segregation "is an essential characteristic of modern (western) societies and allows for different kinds of social relationships to be established and maintained". [Rac75] If everyone has access to all information related to an individual all the time, relationships would no longer be possible. Figure 1 shows an example of an identity that contains several partial identities.

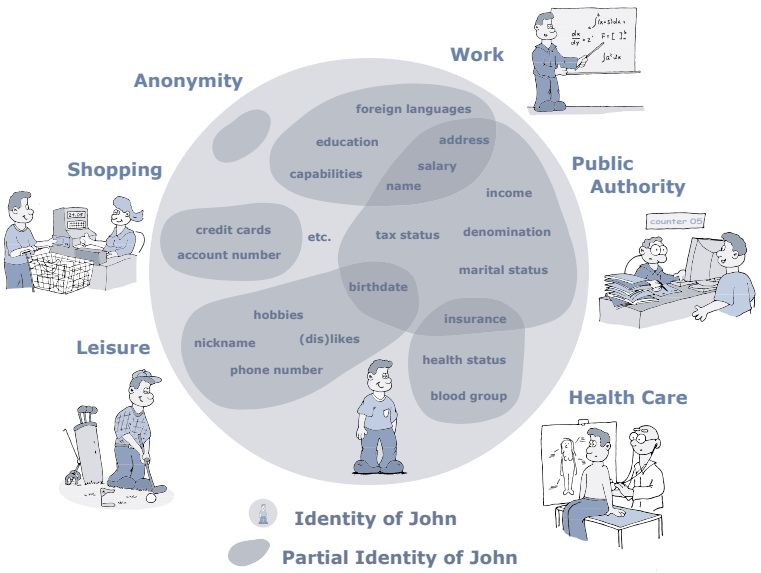

Figure 1: An identity comprised of multiple different identities.

Areas of life. Contexts can be grouped into areas of life as shown in Figure 1. Areas of life are sufficiently distinct domains of social interactions that fulfil a particular purpose (for the data subject) or function (for society). Areas of life are thus defined mainly by the relation of an individual to the society.

Digital personae. The establishment and maintenance of relations takes place offline as well as online. In the online context, representations of individuals (partial identities)

can be referred to as *digital personae*[1] or digital partial identities. It should be mentioned here, that the notion of Personas is known as an important utility within the contexts of application design. The basic idea of those Personas has been founded by Alan Cooper[2] in 1983. It refers to creating patterns of human beings by determining common characteristics of how users (would) utilize software applications. Personas in the sense of application design enable developers to shape their applications in such a way that they address particular requirements of their users projected onto such a Persona. So, it is clearly differing from digital personae this section is taking into account.

Digital personae are (online) representations of individual's partial identities. This is, however, still a vague notion that needs further explanation. For this paper the starting point will be the definition of digital personae given by Roger Clarke: "The digital persona is a model of an individual's public personality based on data and maintained by transactions, and intended for use as a proxy for the individual."[3] This definition clearly reflects the issue of representation. Furthermore, Clarke makes a distinction between projected digital personae and imposed digital personae. *A projected digital persona* is created by the individual and is strictly related to the way this individual wants to present himself. A MySpace profile page is a good example of this form. The individual has significant control over the image created by the audience. Users of Social Network Sites (SNS), of which MySpace is a well-known example, take great pains to construct and foster a certain image of their identity by means of typography, images, language, links, preferences, etc.

In contrast, an *imposed digital persona* is created by institutions based on the information they collect(ed) about an individual, and this persona has a certain function related to their task. Part of such a persona might be that Peter is unemployable because of his handicap, or that he is lonely and terminally ill. These images of his identity are likely not to be those that he himself would like to project to the world, but are rather the image created by the outside world and associated to him.

Recent examples in the Netherlands are the Personal Internet Page (Persoonlijke Internet Pagina, PIP) or the Electronic Child Database (Elektronisch Kind Dossier, EKD). However there are much older examples of imposed digital persona that are used since many decades such as estimating an individuals' creditworthiness, e.g. the Schufa in Germany.

Both projected and imposed personae have effects on the individual. People may find Helma a cool girl because of her MySpace profile, whereas her mother may judge her to be dull. Peter's environment will behave according to the persona imposed upon him by the various institutions. Based on digital representations, decisions are made, some of which are unknown to the affected individuals. However, the decisions clearly have an influence on these persons.

With regard to the projected persona and the imposed persona Clarke states: "The individual has some degree of control over a *projected persona*, but it is harder to influence

[1]The term personae is the plural form of persona. Some authors use personas as plural, however, we prefer the Latin form. Thus, the term *personas* means the same as *personae*.

[2]See http://www.cooper.com/journal/2003/08/the_origin_of_personas.html (last visited: February 2011)

[3]See: http://www.anu.edu.au/people/Roger.Clarke/DV/DigPersona.html (last visited: February 2011)

imposed personae created by others. Each observer is likely to gather a different set of data about each individual they deal with, and hence to have a different gestalt impression of that person."[4]

The amount of data collected and stored about individuals is only growing. This is due to the difficulty, or impossibility even, to erase digital data. Once disclosed on the Internet, information will never again become private. This phenomenon contributes to the risk of collapsing contexts, i.e. separate contexts are connected or combined, when digital personae representing an individual are connected.

Lifespan. The lifespan of a human being is the range of time from the emergence of the first information that is related to this specific human being otherwise legally known as the data subject (a time period from the moment of birth until death or even thereafter) until the point in time when no more personal data is generated. Here, the verb 'generate' refers to new information becoming available to other persons than the former data subject. Hence, lifespan refers to the temporary aspects of privacy and identity-management and in particular to the challenges involved in realising (privacy-related) protection goals over very long periods of time. This aspect closely corresponds to the claim to cover identity management "from birth till death". Without going into unnecessary detail on ethical and philosophical questions about what constitutes human life, the lifespan broadly covers the time from the first diagnosis of a pregnancy until long after the data subject's death. This is so because often times the estate of the dead reveals information about them. According to the Privacy Directive [Eur95], only data referring to a (living) natural person is considered "personal data", Art. 2. However, an individual may want to control how information concerning him will be treated after his death. With this definition, most lifespans will never end in theory (because one can never be sure that no more information will be found). But in practice one can consider an "information lifespan" over when the probability that such information will appear and, can univocally be attributed to the deceased individual becomes negligibly small. Another issue to take into account is that data concerning deceased people can contain information that is relevant for, or refers to, others, such as genetic data.

2.1.2 Identity in Formal Settings

In this section we describe the lifecycle of (partial) identities. Partial identities of an individual differ from identities in that they are not necessarily used to "sufficiently identify this individual within any set of persons"[5]. So, they are qualified for managing one's privacy.

There are a number of events related to the evolvement of the identity which can be described as different phases [HPS08]:

- *Establishing a partial identity* means that the partial identity is created by or assigned to a person.
- *Evolving a partial identity* includes

[4] See: http://www.anu.edu.au/people/Roger.Clarke/DV/DigPersona.html (last visited: February 2011)

[5] Definition of *identity* as given in [PH10]

  - the usage of the partial identity by the holder
  - the usage of a partial identity by others. Their maintenance includes observing or storing it and possibly by applying all kinds of data processing operations.

- *Termination of a partial identity* means deletion or suspension of the partial identity. Note that in some specific cases it can be possible to re-establish suspended partial identities.

All phases are relevant for formation of partial identities.

Identifiers. Personal identifiers are alpha-numeric strings that can unambiguosly be linked to a certain person. Such a personal identifier may be created for the whole lifetime or even beyond (e.g. in Germany a number created for the pension insurance fund that may also pay to an insurant's wife).

"All [EU] countries use general identifiers that are not restricted to use within one specific application or sector. Such identifiers would in principle be more suitable for identification purposes than sector/application specific sectors, since they are less likely to be restricted to a limited user group. However, in some countries their use is restricted by law, precisely in order to avoid that governments can link personal data about a specific person across different sectors, which is considered to be a privacy threat in some countries. This can render them unusable for cross border authentication purposes." [IDA07, p.36]

Formal Identities. The establishment and use of formal identities usually takes place by institutions. They create an identity or identifier on the basis of a legal obligation. Data about individuals related to the specific context or purpose of the identifier is connected to the identifier. All together, the sets of data form partial identities.

Our general formal identity is given or created by the state. When a child is born, the parents have to register the child at the governmental institution of the place of birth. When the child is registered, the government provides a formal identity in the sense that there is a record of birth made up. This record contains the name(s) of the child, date of birth, place of birth, and information about the parents. The child will also receive some unique identifiers, usually numbers. For instance, in Germany the number of the birth certificate is one identifier, and the newborn is also assigned a unique number for tax purposes whilst in the Netherlands, the newborn receives a BSN[6], which is used in multiple public sector contexts.

The name(s) of a child are chosen by its parents, but formally it is often the state that assigns the name to the child and therefore it is the state which creates the newborn's identity in the formal sense. There are restrictions on first names to be proposed for the newborn. Some trade marks or sensitive names (from a historical perspective or because they are immoral) will be refused by the authorities. Famous in this respect is the French case regarding the parents who wanted to name their little girl Mégane Renauld pronounced the same as Renault Mégane, a popular French car at the time.

[6]BSN – BurgerServiceNummer (engl.: Citizen Service Number)

Although the courts ultimately decided not to overrule the parents, they could have done so.[7]

With regard to the family name, the child receives the name of its father or/and mother (in the Netherlands at least the parents can choose which family name their hild receives). Married parents in Germany either already have settled for a family name when they married, which automatically transfers to their children, or the parents have to select one of theirs to transfer when the first child is born. The chosen family name is given to all following children. If the parents are not married, the name of the mother is given by default if the mother does not declare that she wants the name of the father to be given. An interesting complication arises when the unmarried couple decides to marry after the child's birth and decide to adopt the father's surname as the family name, because then the child's surname will change as well. Also more complicated naming schemes exist. In Spain, for instance, children receive both their mother's and father's surname and hence have a double family name. In Ireland the parents decide on the family name of the child when registering the birth and may change this at a later stage; there is no restriction on using composed family names. Not only names of children may change over time. In many countries it is customary or even a legal obligation that married women acquire their husbands name when they marry. Also individuals may request a formal name change, due to, for example, harassment, cultural issues (for instance, in the US many immigrants have requested name changes to better blend into the US culture [Sca96]), or witness protection schemes. In other words, names are not particularly stable identifiers for individuals, which is one of the reasons for the popularity of numbers as identifiers in formal contexts.

The unique identifiers (the numbers) given will, usually, be used throughout the individual's entire life in interactions with the government. These interactions include for instance taxes and subsidies as well as the distribution of travel documents (passport) or identity cards and driving licenses.

The information will be kept in the official registers: "data collections held and maintained by public authorities, in which the identity attributes of a clearly defined subset of entities is managed, and to which a particular legal or factual trust is attached (i.e. which are generally assumed to be correct)."[8]. The identity information can be kept in municipal administrations, local records, as well as at a central governmental level. If a person moves from one city to another city, he generally has to deregister in his old hometown and register in the new one.

After decease, a death certificate is created and the death is registered in the local records. The data remain archived for, amongst others, genealogy and statistic purposes.

2.1.3 Formal Identities in Different Contexts

Next to the general formal identity as described above, a number of context-related partial identities are created during the lifetime of the individual. This section describes these identities in four key areas of life, namely government, education, healthcare, and

[7] See [Whi04]: "The court's opinion emphasized that the parents had not any 'arrières-pensées' – that is, any unacknowledged or ulterior intentions, and that the car model in question would likely go out of production by the time the child reached school age."

[8] Definition token from [IDA07]

employment.

Government. Soon after the birth of an individual, the government grants a birth certificate and thus creates the identity of the individual for governmental registries. Probably, the certificate also contains a number or other identifier which is then connected to the individual. From then on, identification of the individual takes place on the basis of this number. Next to interactions between the government and the individual, other interactions may use the same identifying number. Many interactions with the government leave traces in the individual's records.

Termination of the identity takes place after death. However, this only counts for the identifier, in as far as the number will be *decommissioned* and will be placed on a revocation list. The records remain, together with the registries.

City administration records also contain information on the date of birth of an individual and its marital status. In tax filings, this information is combined with information on income and some insurance. Usually, tax filings use the same identifier as provided by the government at birth.

Once an individual dies, the information is used to identify the heirs and to get all administrations correct.

As described above, the government creates a general formal identity for each individual. However, next to this general identity there may be many partial identities, related to specific contexts. These identities can be separated, but may be connected via the general identifiers of the individual. In the governmental domain driving licenses, travel documents, taxes and subsidies were already mentioned as specific contexts. These smaller contexts all have their own identity information concerning the individual. Other examples are marriage, changes in family situation and permits for building or parking.

Education. Another important context where a partial identity is used is the educational domain. In principle, all individuals go to school at some point in time and many go to kindergarten before entering a school career. In kindergarten, as well as in school, records are created on the (social) development of the child.

Once an individual starts visiting school an identity will be created by the school. Probably, only name and address details together with date of birth are used to directly identify a person, whereas additional data on personal development give a more profound view of the individual. However, it is more likely that the educational institution also creates an identifying number which is used to indicate an individual. During the educational life-cycle, data about grades and certificates, personal comments from teachers, and general observation data are added to the records, thereby shaping the pupil's or student's identity. Most educational institutions use electronic systems with pre-fixed tables and schemes to describe the development of the child. Not only skills such as writing and counting are included, but also social skills such as "How does the child react to the teacher/strangers?"; "Can the child play/work on his own?"; "Does the child have many friends?" etc.

Usually personal data on the individual and possibly his relatives (e.g. parents, brothers or sisters) are disclosed to the school, and the way how to prove the authorisation for attending the courses is communicated – e.g., by simply stating one's name or by

showing an assigned chipcard. As soon as these partial identities are created and the individual himself begins to use them like by attending school, he begins to further develop those partial identities – and thereby also to manage identity – himself.

Files regarding education will contain personal info and grades as well as an overview of which education someone follows. When the age of the individual and his/her educational level are rising, files will also contain information about financial support and whether a student is living with his/her parents or not.

Occasionally, data will be shared amongst different educational institutions, for instance when someone switches to another school or goes from secondary education to a university. At least diplomas will be needed, but probably also grade lists and other information. This information might be exchanged either directly between the different educational institutions or the individual gets a certificate from the first institution that he shows to the second institution.

When an individual finally finishes education, the partial identity could be terminated. However, diploma or certificate information remains stored in order to be able to verify the authenticity, implying that the identity is maintained and remains.

Student records give an insight in the number of students and the kind of students someone is studying with and they have an index of registered certificates. The registration of these certificates can be shared with other instances than the school itself, for instance when there is a verification needed.

It is also possible that schools collect information on extra curricular activities of their students.

Health Care. Prior to the newborn's birth, data will be collected from the mother-to-be and the pregnancy that will become part of the newborn's identity. Peculiarities during the pregnancy and certain developmental or genetic defects will be recorded and become part of the medical record that is created at the child's birth. Before and after birth, general practitioners, specialists, hospitals and other health care professionals exchange patient and medical information. Some of this information will also be shared with health insurance companies (think of treatment bills) in order to be able to conclude insurance policies. Medical data may also be collected by research institutes and government agencies for epidemiological surveys. In these cases the data usually will be anonymised.

From an early stage, records are kept on vaccination and blood group. Depending on the events that occur during someone's life, extensive medical records may develop. Furthermore, Health care during someone's life can include somatic health care as well as mental health care.

Employment. In order to get employed, people need to have a social security number (provided by the government or tax services). Employers will create a file which includes information on name and address, educational level, kind of work, a complete CV, the bank account, and salaries or wages. Probably, the employee also gets an employee number from his company.

Capabilities will be tested and there is information about the function and status of an individual (employee/employer, freelancer, etc.).

Identity management related to employment includes both the situation of being employed and being unemployed. Once an individual becomes unemployed, he may apply for social security and will probably be registered as job-seeker. There can be a duty to apply for jobs, which is supervised by the government.

2.1.4 Identities and Social Networks

Typically people do not live alone and independent for the whole of their life; they start with parents, some will marry and have children and grandchildren. Usually many other relatives exist; ones they know about, others they are nor aware of. Most people also have a number of friends, schoolmates, and colleagues during their life. Although schoolmates are people one gets to know at school and colleagues are people one gets to know at work, usually the link to them can not be described formally. The social network people form and live in nevertheless affects their privacy as much or even more than the formal areas described above. This holds even more nowadays in the time of Web 2.0 because many people transfer their real social network to social networking software and begin to form new social networks on the Internet. Often they are not aware of the fact that the people they address with postings on a web site are not only friends, but often include every user on the world with Internet access.

In this document, we focus on the following aspects of social networks that affect someone's privacy (also in the formal areas) and may result in persons with the inability to protect themselves against privacy breaches:

- Data belonging to more than one person
- Data about other persons
- Data about dead persons

2.2 Fundamental Definitions within Privacy Throughout Life

2.2.1 General Definitions

Most of the common definitions are derived from the Data Protection Directive [Eur95], from the ePrivacy Directive [Eur02] as well as from previous work in workpackage WP1.3 as follows:

Data subject. An identifiable natural person[9], which is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity [Eur95, Art. 2a].

Data subject's consent. Any freely given specific and informed indication of his wishes by which the data subject signifies his agreement to personal data relating to him being processed [Eur95, Art. 2c].

[9] We also use the term "individual" or "human being" for "natural person".

Data controller. The natural or legal person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing of personal data; where the purposes and means of processing are determined by National or Community laws or regulations, the controller or the specific criteria for his nomination may be designated by national or Community law [Eur95, Art. 2d].

Processing (of personal data). Any operation or set of operations which is performed upon personal data, whether or not by automatic means, such as collection, recording, organisation, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction. This also includes the action of anonymisation or pseudonymisation of personal data, even if after such action the data may no longer constitute personal data [Eur95, Art. 2b].

Privacy-relevant data processing. Not only processing of personal data may affect the privacy of an individual. For instance the provision of ICT systems which enable linkage of data can be relevant to the private sphere of the individual because this linkage may yield personal profiles on which decisions are based [HM07, RBB+08]. Similarly, ICT systems which aggregate data to group profiles instead of personal profiles may affect the private sphere of each individual concerned by enabling her discrimination [Phi04]. Further, not all parts of an ICT system that processes personal data touch those data themselves; still they can be relevant for the system's decision-making based on individuals. Note that with service-oriented architecture this phenomenon is by no means rare, but prompts questions to the responsibility for data protection of the data subjects concerned. The term "privacy-relevant data processing" encompasses all these ways of data processing.

Data processor. A natural or legal person, public authority, agency or any other body which processes personal data on behalf of the controller [Eur95, Art. 2e].

User. User means any natural person using a publicly available electronic communications service, without necessarily having subscribed to this service [Eur02, Art. 2a].

Developer of an ICT system (or system developer). A natural or legal person that is involved in conceptualising, designing and/or implementing an ICT system. Taking a wide view on the term "system", "system developers" are meant to include "application designers".

Application provider (or service provider). A natural or legal person that operates an application based on an ICT system and offers it to users.

Policy maker. A natural or legal person with power to influence or determine policies and practices at an international, national, regional, or local level. This comprises law makers, standardisation organisations for technical standards, and supervisory authorities. In addition privacy organisations which are not institutionalised by a State can play

a role as well as media such as the press or bloggers – these can be considered influential to policies although the narrow term of "policy maker" usually does not comprise media.

Caretaker. A natural or legal person with some responsibility for an individual, for example, a parent, a teacher, a trainer or an employer. It is sufficient if the person feels the responsibility. In the area of privacy, a caretaker should try to empower others in self-determination.

Stage of life. A stage of life of an individual with respect to managing her privacy is a period of life in which her ability to do so remains between defined boundaries characterising this stage of life [CHP+09]. Every individual during her lifetime passes through one or more stages during which she does not have the ability to understand the consequences of data processing relevant to her private sphere or to act upon that appropriately.

Delegation. Delegation is a process whereby a *proxy* (also called delegatee or agent) is authorised to act on behalf of a *principal* (also called delegator) via a mandate, i.e., transferred duties, rights and the required authority, from the principal to the proxy. The field of delegation has been discussed by various authors, mainly aiming at technical solutions for specific scenarios. Putting the focus on privacy aspects, we deviate a bit from the definitions used in [PRMD10] or [Cri99]. In our setting, both principal and proxy are natural persons.[10] The delegation may be invoked by the principal herself, but there are also cases where other entities explicitly decide on the delegation (for example, in the case of incapacitation of person the guardianship court) or where the delegation is foreseen in law (for example, when parents are the default proxies of their young children). The power of proxy is usually assigned for a specific period of time.

Data handling policies. Data handling policies were already defined within PrimeLife as a set of rules stating how a piece of personal data should be treated (see[ABB+09] for details).

2.2.2 Data Types

During one's lifetime many different kinds of data appear and many different data may be disclosed by the data subject. This might be data about the data subject herself or data about others. The following data types can be defined:

Personal data. Any information related to an identified or identifiable natural person. Natural persons are only living individuals but neither deceased nor legal persons [Eur95, Art. 2a]. Note that [Art07] refines this definition by elaborating on "any information", "relates to", "identified or identifiable" and "natural person". This work is quite helpful for practitioners; however, there are still open issues, in particular concerning new

[10] It is also possible that legal persons become proxy, for example, organisations for children's welfare, public youth welfare office. And under certain circumstances even the principal may be a legal person. However, broadening the view to legal entities overstrains the scope of this text and may be a task for future research.

technologies and concerning intercultural settings where the terms may be interpreted differently, for example, pointed out in [RBB+08].

Special categories of data[11] /"sensitive data":

- Personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and data concerning health or sex life (these categories of data are also referred to collectively as "sensitive data").
- Personal data relating to offences, criminal convictions or security measures.
- National identification numbers or any other identifiers of general application.

Note that the sensitiveness of data perceived by an individual may be different from what is expressed by the special categories according to Art. 8 of the European Data Protection Directive [Eur95, Art. 8]. Moreover, concerning long-term risks in an unpredictable setting, the view on the sensitivity of an individual's data should be broadened, as proposed in [CHP+09] based on [HM07]:

- "*Data may be static, or changes are quite accurately predictable:* Data which are static over time and are disclosed in different situations enable linkage of related data. Examples for static data are date and place of birth. Similar to static data are those which are quite accurately predictable or guessable because they follow some rules. [...] If static identity information is being used for purposes such as authentication, this bears a risk because these data cannot easily be revoked and substituted [...].
- *Data may be (initially) determined by others:* Data which the individual concerned cannot determine himself (for example, the first name) may persist or it may take a significant amount of time or great effort to change them. A special case is the inheritance of properties from others, for example, the DNA being inherited from the natural parents.
- *Change of data by oneself may be impossible or hard to achieve:* If data are static (see above) or if data are not under the individual's control, wilful changes may not be possible. Examples are data processed in an organisation.
- *Inclusion of non-detachable information:* Data that cannot be disclosed without simultaneously disclosing some side information tied to the data should be prevented or the individual should at least be made aware of this. Examples are simple sequence numbers for identity cards which often reveal sex, birth data and at least a rough timeframe of when the identity card was issued [HM07].
- *Singularising:* If data enable to recognise an individual within a larger group of individuals, the individual privacy may be invaded by tracking or locating, even if other personal data of the individual are kept private.
- *Prone to discrimination or social sorting:* There are no data which are definitely resistant against a possible discrimination forever. This does not need the individual to be identified or singularised. If some people disclose a property and others

[11] See [Eur95, Art. 8]

> resist to do so, this already allows for social sorting or positive discrimination." [CHP+09]

Partial identities. Personal data can be represented by so-called digital identities consisting of attributes, i.e., sets of personal data. A (digital) partial identity is a subset of these attributes – depending on the situation and the context both in the physical and digital worlds – that represents an individual [PH10]. Note that a digital identity usually is only growing, never shrinking over time because it is very hard – if not impossible – to erase widely used digital data [HPS08]. Consequently, it cannot be expected that privacy-related activities, such as disclosure of personal data, or their consequences are revocable.

2.2.3 Areas of Life

Individuals interact with other individuals and organisations in many different relations, all of which are connected to different roles of the individual. Identity was already defined by Goffman as "the result of publicly validated performances, the sum of all roles played by the individual, rather than some innate quality". [Gof59]

The data set which characterises a role can be regarded as a partial identity. Depending on the context (relation) between the individual and the person or entity they interact with, certain information is disclosed or not. The information disclosed and characteristics associated to the individual are attributes of this individual. Individuals from a data perspective can therefore be seen as a (large) collection of attributes. For a concrete partial identity the attributes take specific values. So 'first name' is an attribute label while 'Peter' is an attribute value.

In daily life, people are subject to various subscriptions and therefore have special behaviours and follow special rules depending on the contexts. They even want to present different faces of themselves, depending on the impression they want to conciliate. Therefore the data subject also distinguishes which audience is allowed to see which data of him/her. Audience segregation is a device for protecting fostered impressions. If everyone had access to all information related to an individual all the time, relationships would no longer be possible.

2.2.4 Digital Footprint

The term "digital footprint" in this deliverable refers to the definition developed in PrimeLife work package 1. Individuals engage in social and economic life and during their lifetime act in many different areas of interaction, such as worklife, leisure, financial services, healthcare, or governmental services. Every person leaves an enormous amount of digital traces during her lifetime. Each action or transaction that is electronically performed or supported provides an information log. For instance shopping and paying with a bank card or credit card, all Internet actions (browsing, click trail), electronic toll systems, etc. The thousands of data together form a digital footprint of the individual. The data contained in the digital footprint can be created by the concerned individual herself, for instance in the above mentioned transactions or when someone creates a profile page on a social networking service (SNS), or the data can be created by others, such

as governmental bodies or businesses.

Furthermore, these data contain partial identities in different areas of life as shown in Figure 1.

Digital footprints are personal data of a person that accumulates in information systems. Most people are unaware of this information and the specific type of information that may be available online. It is also a matter of awareness to get digital footprints visible and inform the user about personal data stored in the web (or in databases). As stated in previous PrimeLife deliverables, ideally only the concerned individual herself should be able to access her digital footprint. The PrimeLife prototype ideas "Show my digital footprint", "Remove my Digital Footprint" and "Central Data Handling Repository" try to realise a first approximation of such a service (cf. Section 4.1).

This chapter shows how digital footprints (personal data of a person that accumulates in information systems or in databases) of persons may appear and develop within someone's life and relates them to lifelong requirements. It is important that persons get legal and technical opportunities to control their digital footprints, for example, by deleting parts of them or by encrypting parts of the digital footprint. It should be noted that probably most of the data in one's digital footprint qualify as personal data because of their context or the combination with other data in a data set, which makes it possible for the data to be indirectly linked to an individual.

2.3 Conclusion

It is useful to see identity not as a single concept, but rather in the respect of individuals having multiple partial identities that literally come into play in different contexts. We have adopted Roger Clarke's notion of digital persona in this deliverable as the digital representation of an identity. It is useful to distinguish between projected personae and imposed personae. The projected persona is how the individual aims to present himself to the outside world whereas imposed persona relates to the image that others create of an individual.

In order to shed some light on differences in the treatment of individuals and to provide a first glimpse of whether partial identities really exist in the real world or whether governments and enterprises create and use a single (holistic) digital identity of the individual, we have explored four specific contexts. The analysis started from a common background for all individuals, the state-created general formal identity which unsurprisingly plays a central role in the context of citizen-government relations.

This chapter showed a number of problems that occur when it comes to throughout life aspects and identity management. The issues can be diverged into three categories, namely; data linking different persons; data about other persons, and; data about dead persons. Data remain available after decease. But already during lifetime, several problems occur because of the increasing (electronic) data exchange and processing, and because data can ever more often be related to more than one person. Also, control is a specific issue. In the case of minors or elderly, control over data can be delegated to others, by law or on a voluntary basis.

Having defined the basic terms, the next chapter will describe the requirements and concepts for privacy-enhancing daily life.

Chapter 3

Requirements and Concepts for Privacy-Enhancing Daily Life

This chapter recalls the objective of data protection and privacy regulation in terms of high-level requirements for privacy throughout life. The chapter refers to legal provisions that regulate these objectives and derives high-level requirements. These requirements focus on general principles which describe what should happen with privacy-relevant data and what should not happen with these data. The following section will seize upon these general principles by adapting them to more specific scenarios or perspectives to derive further requirements.

3.1 High-Level Requirements for Privacy Throughout Life

In this section, high-level requirements regarding transparency, data minimisation, fair use, data subject's identity management as well as change management are analysed. But also the high-level requirements regarding practicability of mechanisms and data handling policies are discussed to help to prevent further risks because of mistakes in data processing and on exercising one's rights.

High-level requirements are derived from changes in society, law and technology. This relates to the implementation of data protection management systems by data controllers to ensure legal compliance and the state of the art in ICT security over time or the reaction to social changes with regard to legal and technical aspects. Societal changes also need to be considered with regard to legal and technical aspects. They have to be recognised and appropriate technologies or legal regulations have to be taken into consideration. Furthermore the assessment of technology and regulations may guarantee a kind of quality assurance.

For the processing and handling of personal data some general characteristics and requirements can be derived from the European Data Protection Directive 95/46/EC [Eur95] as well as the OECD Guidelines on the protection of Privacy and Transborder Flows of Personal data [OEC80].

If privacy has to be considered over a long period of time, some problems will emerge:

- *Technical:* Proclaiming that a certain cryptographic technique will be good enough for 40 years or more, is considered to be ridiculous.
- *Legal/sociological/political:* In a time of 40 years or more, laws, regimes and structure (i.e., common ideas) of society can change drastically (cf. [SA08]). What can be regulated by law, politics, and social pressure, might change.
- *Societal:* The concept of privacy, i.e., what is considered to be private or sensitive, might change over time. This implies that revocability of techniques might also be necessary.

In a long-term setting there surely will be some dynamics in policy: both the policy of society at a larger scale and the quite individual policy of a human being in relation with interaction partners [CHP+09]. This poses challenges for technological solutions, in particular:

- Which aspects of technology, which rules implemented in technology need to be addressable by such dynamic changes?
- Which aspects must not be changeable, thus allowing the individual to trust that her expectations will be met, no matter what?
- What are the abusive potentials of new technologies, if not used in a way that one had in mind in the first place?

The starting point of the elaborated high-level requirements is the situation of today: There appears to be at least a common basic understanding of privacy and a consensus that the current baseline will never change, at least in democratic societal models. However, all solutions will have to cope with upcoming changes and cannot – and should not – freeze the status of today.

3.1.1 Openness, Transparency, Notice, Awareness, Understanding

Transparency is one of the general principles in our society and also with respect to privacy. It is a necessary principle for estimating privacy risks and decision making concerning privacy-relevant issues, and it is also a prerequisite for further action such as asking all data recipients for access to one's personal data or requesting their erasure. Thus, it is one of the main principles with regard to the data subject's rights, and many requirements within this text will refer to transparency. As it is stated in the European Data Protection Directive [Eur95, Art. 6] that Member States shall provide that personal data must be processed fairly and lawfully, which also means that the data subject must be informed about all data collected and processed about him. Therefore transparency has to be ensured with regard to data processing (data flow, data location, ways of transmission, etc.) in respect of users of the product or service as well as data subjects. An informative, up-to-date and understandable, well-searchable description of the product or service has to be provided to the user (who has to get simple access to those provisions). The data subject has to be informed to whom data are further processed.

Transp-Req a): For all parties involved in privacy-relevant data processing, it is necessary that they have clarity on the legal, technical, and organisational conditions setting the scope for this processing (for example, clarity on regulation such as laws, contracts, or privacy policies, on used technologies, on organisational processes and responsibilities, on data flow, data location, ways of transmission, further data recipients, and on potential risks to privacy).

The right to informational self-determination furthermore includes the right to know, who knows what about the data subject [Eur95, Art. 15]. With regard to this, [Eur95, Art. 12] furthermore states, that the data subject has the right to obtain from the controller knowledge of the logic involved in any automatic processing of data concerning him.

Awareness. The requirement transparency is very much related to awareness. First of all data subjects have to be aware of the identities that are created by them in daily life or in the web. The requirement awareness is of special interest, when identities of individuals/users are created about them by others. In these cases individuals are not aware about the existence of formal identities and they do not have any control. Therefore all parties involved in privacy-relevant data processing, in particular data subjects, should be made aware of potential risks to privacy and ways to deal with these risks, for example, in privacy policies. But it has to be taken into account that too much information may overwhelm the data subject and in this case awareness is also not given any more because the data subject can not use the information properly. Creating awareness also means to find a balance of appropriate information of the data subject. Furthermore, the expectations on awareness may vary in different societies (to be more specific, this refers even to different cultures and subcultures). As society is changing, also the responsibilities may change. But in conclusion the focus should always lie on the data subject. The question on when or who will be informed by whom and how has to be clarified.

Transp-Req b): Schools or education centres should make individuals aware of potential risks to privacy and ways to deal with these risks.

Transp-Req c): Data controllers and data processors should make their employees aware of potential risks to privacy concerning data processing and ways to deal with these risks.

Transp-Req d): Parents should make their children aware of potential risks to privacy and ways to deal with these risks.

Transparency of What Is Irrevocable and What Is Revocable. In a lifelong context certain information on data subjects needs to be revocable, whereas other information might not be. Legal provisions acknowledge such revocability for personal data,

by the right to have data deleted, and also in copyright law, when the author decides so. Therefore the Directive states, that every reasonable step must be taken to ensure that data which are inaccurate or incomplete, having regard to the purposes for which they were collected or for which they are further processed, are erased or rectified [Eur95, Art. 6 and 12]. The data controller needs to erase or block data that do not comply with the provisions of the Data Protection Directive, in particular because of the incomplete or inaccurate nature of the data.

However, in the case of copyright law, already published material cannot be recalled. The right of the author merely allows stopping further publishing. The material also will be available in archives that already acquired it. In an information society, this might mean that the information is still going to be widely accessible. Clear rules need to be defined that make transparent if and under what circumstances information will be available forever.

> *Transp-Req e):* For all parties involved in privacy-relevant data processing, it should be clear under which circumstances decisions are revocable/irrevocable and what the potential impact can be. In particular, data controllers should inform data subjects on to which degree their decisions (such as consent to processing of personal data or distribution of these data) are revocable or not.

Transparency and Accountability. The law in some cases already provides provisions, which oblige data controllers and processors to log their processing of data. These logs need only to be kept for a certain period of time, and then need to be deleted.

As log files may contain personal data about the data subject and those who are processing the data, it has to be considered under the lifelong perspective a historical dimension, to allow later access for research purposes and thus data may need to be stored in archives. Sufficient logging mechanisms need to be implemented. There also has to be sufficient information of the data subject what kind of data are stored within a log and for how long the log is accessible under which conditions (for example, within the privacy policy).

> *Transp-Req f):* Data controllers and data processors should keep audit trails on the privacy-relevant data processing.

> *Transp-Req g):* For audit trails, data controllers and data processors have to define and make transparent (at least within the organisation and for supervisory authorities) which information is logged for how long.

> *Transp-Req h):* For audit trails, data controllers and data processors have to define and make transparent (at least within the organisation and for supervisory authorities) who can get access to the log data under which conditions.

Note that there may be the need of a secondary audit trail to log all accesses to the primary audit trail if it contains privacy-relevant data. Of course this cannot be infinitely repeated in a recursive process by introducing a third, forth etc. audit trail, but instead controlling the access of an audit trail may be realised by applying the four (or more)-eye-principle without the possibility of one party to access the data on its own. Also, audit trails should be designed in a data minimising way, e.g., by using pseudonyms so that the log file can be analysed in a first step without directly identifying persons, but offering a second step, e.g., in the suspected case of misuse, where more personal information is provided.

Transparency of the Logic Behind Privacy-Relevant Data Processing. The data subject has a right to know who knows what about him or her. Therefore, the logic behind the processing, especially the processing of personal data with regard to profiling has to be described in detail to guarantee transparency for the data subject. The right of information therefore comprises not everything that is technically possible, but the processing of personal data, which is actually foreseen and controlled by the processor. If personal data are analysed in a statistic-mathematical way, to classify the user by interests or purchasing power, within the constituency, these mechanisms have to be revealed by the processor. Important is the principle of function of the application programme, so the user may understand how the assessment and the classification is derived from his personal data and which relevance the personal information have within the processing system of the processor.

Transp-Req i): Data controllers and data processors should inform data subjects about the logic behind data processing (for example, in profiling systems) in a comprehensible way.

Transp-Req j): In case other regulation inhibits detailed information for data subjects, data controllers and data processors should make the logic behind data processing transparent for supervisory authorities.

Transparency on Linkage and Linkability. During an individual's lifetime considering the development and growth of digital life and interaction the probability of data breaches affecting an individual, and therefore the probability of linkability raises. Furthermore, taking the assumption of Moore's law into account to which microchip complexity doubles every two years, future computational powers will keep increasing exponentially and facilitate linking of data. Therefore data controllers and data processors should make transparent for data subjects, under which conditions personal data will be or actually are linked (for example within privacy policies). This is necessary to make the transferral of data across contexts transparent for the data subject.

Transp-Req k): Data controllers and data processors should make transparent for data subjects, under which conditions (potentially) personal data may be, will be or actually are linked.

Privacy and Security Breach Notification. Data breaches that affect an individual as well as the possibility of linkability need to be prevented. It has to be in the control of the data subject to decide where linkability is allowed or even required. It has to be transparent for data subjects where linkability is possible or already conducted. Therefore data controllers and processors should inform data subjects and supervisory authorities timely on privacy and security breaches and give advice on how to cope with the consequences.

> *Transp-Req l):* Data controllers and data processors should inform data subjects concerned and supervisory authorities timely on privacy and security breaches and give advice on how to cope with the (potential) consequences.

3.1.2 Data Minimization

One of the general principles and one of the high-level requirements that aim at ensuring privacy for life is data minimisation. In general, only a minimum of data, strictly necessary for a particular activity and strictly relating to a purpose of processing, should be processed. Because of the general character this principle appears permanently in several stages of life.

Personal data disclosure should be limited to adequate, relevant and non-excessive data as stated in Art. 6 (1)(c) of the Data Protection Directive [Eur95, Art. 6]. It means that data controllers may only store a minimum of data that is enough to run their services. Implied in this requirement is that data needs to be provided on a need-to-know basis and stored in a need-to-retain basis. This requires the requester to specify the purposes of collection, processing and storing of data. Data should be deleted after the requestor's end as soon as the specified purposes of data collection are met. Data minimisation (incl. prevention of undesired linkage and linkability) in general covers the facets minimal quantity, minimal timeframe and minimal correlation possibilities:

- *Minimal quantity* – limiting disclosure: only disclose those data that are strictly necessary for fulfilling the given task. Data not necessary for the given task should not be disclosed or even retrieved. After fulfilling the particular task necessary data should be erased if there is no legal or consented purpose for further processing.

- *Minimal timeframe* – limiting availability: after usage, data should be discarded. To enforce this, legal, organisational and cryptographic tools can be used. Default retention times after which the data are automatically deleted if not specified otherwise have been proposed, for example, for content on the Internet [MS07].

- *Minimal correlation possibilities* – limiting linkability: advanced data mining technology can allow data controllers to construct links between different partial identities of the same entity. The entity can try to prevent this by running the same data mining technology, upon requests to provide information. This assumes however the same knowledge as the data controllers, which might include invisible links between them (for example, one data controller acting under different pseudonyms).

Data controllers might also try to construct links between partial identities of different entities. From a data subject's point of view, this is very hard to protect against.

DatMin-Req a): Data minimisation means to minimise risks to the misuse of these data. If possible, data controllers, data processors, and system developers should totally avoid or minimise as far as possible the use of (potentially) personal data, conceivably by employing methods for keeping persons anonymous, for rendering persons anonymous ("anonymisation"), or for aliasing ("pseudonymisation"). Observability of persons and their actions as well as linkability of data to a person should be prevented as far as possible. If (potentially) personal data cannot be avoided, they should be erased as early as possible. Policy makers should implement the data minimisation principle in their work, be it in law making or technological standardisation.

Minimal quantity and sensitiveness. To guarantee storage of minimal quantity of personal data, it is absolutely necessary to inform the data subject about personal data stored and in particular about the use of these personal data. Mostly collected data of a data subject are used for profiling and for data mining. Service providers want to offer their service in the best way to their customers to increase the acceptance and, thus to increase their revenues. Therefore, they use profiling based on behavioral targeting. Such customer care mostly also comprises specific offers to a user of the service. Many users do appreciate these offers. But they do not know the data mining behind. There is no transparency about which data are stored and used for profiling and for how long they are stored. In many cases the privacy policy of the service provider does not even mention the fact of profiling or data mining or the customer does not have the chance to use the service and not to be targeted. In conclusion it is necessary that the data subject can decide if he wants to get extra, "personal" offers and therefore is part of the profiling system, or not.

DatMin-Req b): Data controllers and data processors, and system developers should minimise the storage of (potentially) personal and sensitive data as far as possible.

Furthermore, most of the service providers do not sufficiently differentiate data they are collecting. It may happen that many sensitive personal data are collected, processed and used for data mining even if the regulations of the Directive determines special provisions for the collecting and processing of sensitive personal data. Also here the data subject is often not aware that these data are collected, stored, processed and used for data mining.

DatMin-Req c): Supervisory authorities and privacy organisations should support individuals, data controllers and data processors, and system developers to fulfil the principle of data minimisation by giving advice concerning concepts and implementations, pointing to best practices and support research and development in this field. This may be done by employing methods for keeping persons anonymous, for rendering persons anonymous ("anonymisation"), or for aliasing ("pseudonymisation"). Observability of persons and their actions as well as linkability of data to a person should be prevented as far as possible. If (potentially) personal data cannot be avoided, they should be erased as early as possible.

Minimal timeframe. Regarding the lifelong aspect, it is often not adhered to the minimal timeframe for the storage of personal data. The data subject needs to get the control about the timeframe of storing of personal data. Therefore, the timeframe of storage and use of potentially personal data has to be minimised as much as possible and there has to be transparency for the data subject. If there is no legal basis for the use, data should be fully erased.

DatMin-Req d): Data controllers and data processors, and system developers should minimise the timeframe of storage and use of (potentially) personal data as far as possible. After that time, the data should be fully erased. This should comprise temporary files or data which have been distributed to other media or recipients as far as possible.

Minimal disclosure. Disclosure of information constitutes one of the key prerequisites for user control through self-determination, which is a core principle for privacy-enhancing identity management systems. Establishing user control creates satisfactory interactions, human well being, and diverse relations. Other important social aspects of this requirement are: consciousness (individuals have to be aware when their data is processed), comprehension (they need to understand what is actually happening when data is being collected) and consistency (data subjects need to be able to anticipate to the changes of people, preferences, and situations).

DatMin-Req e): Data controllers, data processors as well as individuals should minimise the disclosure of (potentially) personal data as far as possible.

Right of access. Regarding the reference to lifelong aspects, it can be reverted to requirements on access control policies. This requirement demands for an option for access control policies to expire after an amount of time. Access control policies should support conditions and reasoning about time. Time can impact the validity of certain conditions in the policies to be used to support policies that might be valid up to some

time or after some time (for example, embargo on data, data that become public after a given time or data that should be deleted after a given time).

With reference to the general requirement of data minimisation, the policy language should support and encourage minimisation of the amount of personal information that is revealed in order to gain access to a resource. The architecture should definitely not assume that all information about the subject is readily available when the access decision is made. Rather, the list of attributes that need to be revealed, or the predicate that needs to be proved, should be explicitly specified by the server, or perhaps even be the result of a negotiation between the client and the server. The client should then have the option to reveal only those attributes that are strictly necessary. This requirement encourages the basic requirement of data minimisation and helps the data subject to control the digital footprint over lifetime.

It corresponds to Art. 6 of the Directive [Eur95, Art. 6], stating that personal data shall not be kept for longer than necessary for the purposes for which the data was collected. After achieving the purpose for which the data was gathered, it has to be erased or rendered anonymous.

Minimal correlation possibilities – limiting linkability. To protect the data subjects, they should have the possibility to decide whether partial identities can be linked to control her partial identities over lifetime.

> *DatMin-Req f):* Data controllers and data processors, and system developers should minimise linkability and linkage of (potentially) personal data as far as possible.

Furthermore, most of the service providers do not handle context separation. In contrary, most of the contexts are linked to get even more information about the data subject and her behaviour. In general contexts have to be separated regarding the lifelong aspect, especially when connecting them is not necessary for the aim of the formal identity.

> *DatMin-Req g):* Data controllers and data processors, and system developers should minimise multi-purpose or context-spanning use of (potentially) personal data as far as possible. They should provide mechanisms for context separation of these data.

It appears that for instance in the Netherlands, one unique identifier was used in different contexts. The use of such a unique identifier should be prevented. The individual should be able to use a range of identifiers with varying degrees of observability and linkability. This means data subjects must have a choice to operate anonymously, pseudonymously or known. They should also be able to use identities provided by public bodies or enterprises, as well as ones created by themselves, to be able to provide certainty about their identity to other entities and therefore promote accountability when required.

> *DatMin-Req h):* Data controllers and data processors, and system developers should avoid the use of unique identifiers which may be used in different contexts. They should use diverse identifiers where possible.

Furthermore, the requirement regarding anonymous and/or pseudonymous access control is significant. Thereafter a data subject shall have the possibility to access a resource in an anonymous or pseudonymous way. For an anonymous access, the server makes sure that the user fulfils the necessary requirements, for example, "age $>$ 18", while the required attributes allow the user to stay anonymous. This is of course only possible if (1) the required attributes (like "age $>$ 18") are applicable to a big number of people and the data subject can therefore not be identified, and (2) the underlying technology supports proving of the attributes in an anonymous way.

Anonymous and/or pseudonymous access control is important for the data subject's control of his personal digital footprint and for the lifelong data control. It gives the opportunity to access a resource without disclosing personal data and without the assignment of the clickstream and thereby extend the digital footprint.

DatMin-Req i): Data controllers and data processors, and system developers should support anonymous or pseudonymous authorisation and access control of users where possible.

Avoid or limit irrevocable consequences. If within a process it appears, that something may have irrevocable consequences for the privacy of data subjects, it has to be ensured that either the data subject has the choice to decide or these consequences should be minimised in general.

DatMin-Req j): Data controllers and data processors, and system developers should minimise irrevocable consequences concerning the privacy of data subjects.

No forced consents by coupling various services. Some Member States have special regulations regarding the coupling of consent or a general principle of only using data for its original purpose[1]. These regulations are based on the essential content of [Eur95, Art. 2(h)]. Thus, the data subject's consent to data processing must be freely given. In case of data coupling, this consent would not be based on the data subject's free choice. Assuming, the data subject could only access a telemedia service in case of giving her consent to use her data for other purposes such as advertising. If the data subject wants to access this service, there is no choice but giving her consent. Therefore, this consent would not be freely given. Prohibiting data coupling therefore ensures the preserving of the essential content of [Eur95, Art. 2(h)]. Thus, "coupling" of data is prohibited.

[1]For example, the German §12 Abs. 3 TMG, which stipulates that the use of user data for purposes other than providing the service or advertising or passing on the data to another firm requires express consent of the data subject. In doing so, the data controller may not make provision of the search service dependent on the consent to use for other purposes if the user has no other access or reasonable access to such telemedia.

DatMin-Req k): For societally relevant services which may be accessed in an anonymous or pseudonymous way, data controllers and data processors should not make the rendering of services contingent upon the consent of the user to the processing or use of her data for other purposes if other access to these services is, not or not reasonably, provided to the user.

3.1.3 Fair Use – Controllable and Controlled Data Processing

The principle of fair use is mentioned in the Directive [Eur95] as well as the OECD Guidelines [OEC80]. There is no clear definition of "fair use", but the core principles of fair information practice are defined as notice/awareness, choice/consent, access/participation, integrity/security and enforcement/redress. The OECD Guidelines, for example, point out, that there should be limits to the collection of personal data and any such data should be obtained by lawful and fair means and, where appropriate, with the knowledge or consent of the data subject [OEC80, No. 7]. Art. 6 of the Directive demands, that Member States shall provide that personal data must be processed fairly and lawfully [Eur95, Art. 6]. In general for all parties involved in privacy-relevant data processing, the processing should be controllable and controlled. The respective responsibilities must be clear, and accountability of the parties involved for their privacy-relevant actions is important. The data processing should be compliant with the relevant legal and social norms.

Control-Req a): For all parties involved in privacy-relevant data processing, the processing should be controllable and controlled throughout the full lifecycle. It should be compliant with the relevant legal and social norms.

For the controllability of the full lifecycle of a data subject, it is important to notice that others can be individuals as well as companies or institutions. From a legal perspective, this topic is covered in the Lindqvist case; the European Court of Human Rights decided that it is not allowed to publish personal data of others on a website without their consent. The court decision can be applied to all forms of publishing personal data, although the Internet environment will be most relevant. Even though the subject seems legally covered, this remains dependent on awareness of individuals and, for that reason, will mainly remain as an ex-post measure of enforcement. Trying to inform concerned individuals before information about them is published would possibly imply that a database is needed of all individuals and their personal details in order to facilitate this. That is not the most desirable solution.

Other difficulties arise in the context where data can be indicated as (personal) data about more than one person, like for instance relationships or medical data about hereditary diseases. Does the fact that data reveal something about oneself as well as about another individual imply that disclosure is prohibited, unless there is consent of the other data subject? That might be problematic in specific cases where the disclosure of these data might be required in order to enable accurate decision taking, in particular with medical data, even though the disclosure of a hereditary disease to one's children

might conflict with the right not to know. And, of course, gossiping would become prohibited. Solutions may be found in the technical domain. For instance, it may be technically enforced that before posting or publishing content, the consent of all individuals concerned have been obtained.[2]

Consent and its revocation is one of the main issues that influence the digital footprint of the data subject. The data subject's consent as defined in the Directive [Eur95, Art. 7a] is one of the most common legal bases for processing of personal data. The Article 29 Working Party elaborated on the preconditions of a valid consent in its working paper and identified four preconditions: consent must be a clear and unambiguous indication of wishes, consent must be given freely, consent must be specific, and consent must be informed [Art05].

In general, users' data should only be accessible to authorised third parties. These include parties that are legally allowed to access the information (secret service, descendants, doctors), or that have been given consent by the data subject. Given the large time-frame, data subject's consent should be limited in time by default (for example, the consent given by parents for their children is limited until children reach legal age and become autonomous to decide about the consent). In addition it should be made clear what will happen if the person who has consented dies – in some cases this will be equivalent to the withdrawal of the consent, in others the person who died may explicitly want his consent to survive for an additional time-frame (for example, as part of the specific legacy). Moreover, it remains to be defined to which of their data minors are allowed to give consent to others (some of these "rights" might also be attributed to their caretakers). Consent should not only be limited in time, but it should be made clear which parts of the planned (and to be consented) data processing is not revocable and what will happen (how quickly) when the consent is withdrawn. Finally there are also situations that do not allow giving individual consent, for example, if the data subject has no possibility for an autonomous statement (e.g. conscious consent) [Sim06, §4a].

In principle data subjects have the right to withdraw their consent at any time. However, revoking one's consent does not imply that the consequences of data processing can also be revoked: The past cannot be altered; data disclosures cannot be "undone". The revocation comes only into effect for the future and only regarding the data controller the withdrawal of consent is communicated to. In practice, the data controller may already have transferred the data to other parties (this may or may not be legally compliant), or because of a data breach the data may have become known by others. The revocation of consent regarding the primary data controller does not affect these further data disclosures. Also, consequences based on the disclosed and now withdrawn data do not become automatically invalid. This shows that revocation of consent is often a merely theoretic concept.

Purpose Binding. [Eur95, Art. 6b] provides a direct legal framework for data processing. Therefore, personal data must only be processed for specified, explicit and

[2]Note that we do not discuss here individuals in the role of a public figure, i.e., a celebrity or otherwise famous person whose actions are the focus of public interest. For public figures, the public's right to be informed by the press may dominate their right to privacy – at least in those contexts of life which are related to their celebrity.

legitimate purposes. According to the above, personal data should be relevant to their processing and to their extend be necessary for the objectives of their processing [OEC80, Part 2, 8]. Purpose binding therefore defines the requirement of processing and using personal data for the once chosen purpose only. This could be granted in different ways: limiting/prohibiting the use outside the given context, or making the context stick to the data (sticky context should be seen as the meta-data of a sticky policy).

One approach to capture both (context) concepts would be the use of sticky policies. Well-documented sticky policies can provide the necessary transparency and the necessary information to the data subject. They also oblige the processor to handle the data for the purpose defined in the sticky policy only. These purposes are pre-defined and specified, the data subject therefore preserves the information necceary for data handling before processing them as well as the data processor does. Such specifying of purposes beforehand and the subsequent use limitation is called downstream usage [ABB+09]. The purpose limitation (or purpose binding) has central importance for business, since it attempts to set the boundaries within which personal data my be processed, and those within which data collected for one purpose may be used for other purposes [Kun07, 2.89, p. 99]. Necessary prerequisite for the above is the comprehensive information of the data subject before data processing.

In summary, personal data must not be used for further purposes incompatible with the original purposes once they have been properly collected. The data subject has to give her consent for every change of purpose, otherwise this data processing would be inadmissable.

Control-Req b): Data controllers and data processors should restrict the processing of (potentially) personal data to a predefined purpose.

Control-Req c): Data controllers and data processors should be specific in the definition of the respective purposes.

Often the problem arises that purposes are interpreted differently by the controller and the user. To avoid this problem, purposes have to be described and privacy policies to be defined in a clear and understandable way. In many cases, purposes are defined inexplicitly because the controller does not want to be tied to clear purposes to have the opportunity to use data for different purposes. From time to time, the definition of processes may change. This also leads to a change of purpose and therefore a new consent or new legal basis for the processing is necessary. The data subject has to be informed and perhaps a new consent should be given. But it has to be kept in mind, that the processing still needs to be feasible for the data controller and the requirement should not lead to the situation that the data controller needs to ask for consent before every processing.

Accountability. When taking into account that technical development can increase the information value of current data, accountability for data processing becomes a specific point of attention. The protection of data and carefully considering the disclosure and sharing of data are key aspects. If processing of personal databases on a legal basis

or consent, the data processor is also accountable for the processing. Within a company there have to be clear definitions on who is responsible for processing and storage of or access to personal data. Data controllers have the responsibility to adequately protect the data in their systems and the use of personal data is bound to certain legal requirements, such as the requirement of an indicated purpose. When a database is used by the controller, clear concepts regarding deletion or other obligations already need to exist.

Control-Req d): If the data processing is based on consent: Data controllers should limit the data subject's consent in time by default.

Control-Req e): If the data processing is based on consent: Data controllers should ensure that the data subject can withdraw the consent without unexpected impacts on his privacy (because of irrevocable consequences).

Control-Req f): Data controllers and data processors should ensure that the parties processing the data are accountable. This includes the definition and assignment of clear responsibilities.

Control-Req g): Data controllers and data processors should prohibit identity theft, especially in situations which may have privacy-infringing impacts.

Sensitive Data. Within one's whole life, many *sensitive data* are collected and further processed and therefore subject of the fair use high-level requirement. The more personal and especially sensitive data are included within the digital footprint, the more complete is the picture of a person within the web. Sensitive personal data are specially protected within the Directive [Eur95, Art. 8] and the processing of sensitive data is prohibited, except under certain clearly-defined circumstances such as when the data subject has given his explicit consent. Over the lifetime of the data subject, the digital footprint may grow and may be accumulated with sensitive data. This should not happen especially without the data subject's explicit consent or on a legal basis, because it raises the possibility of linkage.

With regard to the general principle of transparency, the data subject should have technical and legal opportunities to control her sensitive data (cf. Section 2.2) that accumulate within information systems or in databases. This may work by claiming the right of erasure by the data subject [Eur95, Art. 12].

Control-Req h): Data controllers and data processors should be extra cautious with (potentially) sensitive data.

To exercise the right of erasure, the data subject needs to know which personal data are stored in which databases or information systems. Therefore digital footprints and especially the processes of collecting, processing and storage of sensitive personal data

have to be transparent for the data subject. This is required to avoid the possibility of linkage of sensitive data. Even if the data subject uses a pseudonym, it should be under her control which data are related to the pseudonym. The data subject needs to have the possibility to decide whether the pseudonym contains too many personal (in this case sensitive) data and maybe in the consequence deletes the whole footprint or even parts of the pseudonym. It should also be under the data subject's control to decide, for example, that data with a certain age are not allowed to be related to the digital footprint (of a person or a pseudonym).

Biometric data can also be defined as sensitive data. The European Union's independent advisory council for data protection issues, the so-called Art. 29 Working Party, comments on the sensitive character of biometric data [Art03, p. 10]. Biometric identifiers are by definition non-revocable, which needs to be considered with regard to the lifelong aspect. Biometric data cannot be changed, once it has been recorded. Static biometric identifiers that do not change over lifetime are particularly critical. If more applications use authentication or profiling based on static biometric identifiers, the risk of unauthorised use of and access to biometric data exists during the remaining lifetime of an individual. Note that more privacy-friendly biometrics are being proposed which prevent re-use of biometrics in a different context, e.g., by only employing dynamic biometric identifiers that will not be released incautiously (like speaking a specific password) or by clever encryption technologies (like biometric encryption[3] or revocable biometrics[4]).

However, often biometric data are translated to code and this code is used to identify or authorise a person. Identification of the person and direct use of the biometric identifier are not necessary in this case. Just like all other digital data, the data can be changed or revoked then. Therefore, transparency of processing of biometric data has to be ensured and there has to be the obligation to notify the data subject in case of loss or stolen biometric data.

Fallback solutions need to be in place addressing a process for individuals who cannot enrol or whose biometric data was compromised. Also fallback solutions for alternative means of access to the service should be provided, especially if biometric solutions are even more widely used and people may be hampered from access to certain services or entering certain locations.

There is also a large number of sensitive data that are not explicitly defined in Art. 8 of the Data Protection Directive. Some Member States have taken the opportunity allowed under Art. 8 of the Data Protection Directive to expand the definition of sensitive data also to data relating to offences, criminal convictions or security measures.[5] There might also be some cultural differences regarding the definitions of sensitive data within the Member States which might lead to the fact, that sensitive data are defined in different cultural ways. Therefore all Member States have to clearly define the definition of sensitive data and have to implement the provisions of the Data Protection Directive in their legislation.

[3] See, e.g., [SRS+98].

[4] See work in the FP7 project "TURBINE – TrUsted Revocable Biometric IdeNtitiEs", http://www.turbine-project.eu/.

[5] For example, Italy (Italian Data Protection Act, Art. 24) or Finland (Personal Data Act, §11).

Organisation of Data Processing. For all data and processes, controllability of full lifecycle is needed: When creating data items or accounts and starting up processes, the deletion of the data items should be anticipated and planned. This is important for data controllers with their professional data processing as well as for individuals who disclose data in a social network or setting up an account somewhere. Not only the existence of data has to be considered, but also its linkability to other data items. This is especially relevant when introducing unique identifiers. Planning the lifecycle also encompasses the definition of procedures for answering user requests (for example making use of right of information) or emergency settings in case of data breaches [Mei09].

Collecting, processing and deletion of personal data within a company have to be defined in detail beforehand and has to cover the full lifecycle of personal data.

With regard to deletion of personal (sensitive) data, controllers and processors have to be aware of the fact that there have to be regulations on deletion when a new database or profile is created or personal data are collected. There have to be clear regulations and processes on when and how personal data have to be deleted if there is no legal basis for processing or no purpose left. There furthermore could be mechanisms to regularly control the legal basis for the processing of personal data and in consequence the deletion if personal data are not necessary any more. In addition controllers and processors need to be aware that conclusion of a contract should not be connected to the data subject's consent in the processing for marketing purposes.

Control-Req i): Data controllers and data processors should conceptualise and plan their privacy-relevant data processing beforehand, thereby covering the full lifecycle of data (from creation to deletion). This comprises to plan the process and set the conditions for potential or factual linkage of data and – if the data processing is based on consent – also for its revocation.

Control-Req j): If identifiers are created, data controllers and data processors should already foresee concepts and procedures for their erasure after the usage period.

Control-Req k): Data controllers and data processors should also plan for emergency situations (for example, privacy and security breaches).

Dealing with Possible Conflicts. With regard to the fair use high-level requirement, there may occur situations with conflicts between different rights of data subjects. For instance, the fundamental right of freedom of speech can prevail the right to privacy as a legitimate interest, also for opinions voiced on the Internet. In German constitutional law a differentiation is made between opinions and facts. Voicing true facts is usually lawful. Voicing opinions is usually lawful, as long as these opinions are not offensive or abusive. A balancing exercise is necessary between the conflicting principles of Art. 8 ECHR (right to privacy) and Art. 10 ECHR (right to freedom of expression) to be performed by courts on a case-to-case basis [ECH10]. Publishing opinions and facts with mostly personal data has fundamental effects to the digital footprint of a data subject.

As voicing opinions is usually lawful, the effect to the digital footprint is also lawful and therefore the data subject can not claim any infringement. But if a balancing of interest is necessary, the data subject may claim against the publishing of data and therefore control his or her digital footprint.

Conflicts of interest furthermore appear when data subjects or controllers are protecting private data while others want to access it. A first remark that can be made here is that this is not necessarily a conflict in which the data subject is involved. Parties that do have legitimate access to the information (for example, hospitals) could refuse access to these data to other parties (national health department trying to control a new pandemic). Even if they would be willing to provide the data because it would serve a good purpose, they might be prohibited by law (or even by their own privacy-preserving technology). In case of emergencies, specific "breaking the glass" policies [Pov00] should be available. In some cases, government/legal actors can intermediate in conflicting interests, using regulations. Automated resolution of conflicts seems to be undesirable [Eur95, Art. 15], but an automated way of notifying users and data controllers that a conflict exists, and technological tools to facilitate negotiation to receive consent seems useful.

Conflicts may also derive from handling "shared data". Some personal data may affect not only one data subject, but several (cf. [Phi04, RBB+08]. Therefore, it has to be clarified how the processing of shared data can be treated by the data subjects concerned. This can be done by technical mechanisms as well as legal solutions, such as clear regulations regarding consent in the processing of shared data.

In the social sphere, i.e., natural persons using Internet applications for personal interests, those persons may publish information not only about themselves, but also about others. Therefore often information about third persons is processed without the consent and even knowledge of the persons concerned. Especially within social networks, photos of friends are uploaded. Even if all people on the picture agree to its publication, they may not like being public some time later. More and more of the contents on the Internet are edited by private persons, through social networking services, such as Facebook, "blogging" or "twittering". It is a legal challenge to clarify the question regarding the enforcement of privacy rights of users that act as data controllers when publishing personal data about others on the Internet and whether this activity is subject to data protection law and what the consequences are [Kor09]. In general, the Directive [Eur95] does not impose the duties of a data controller or an individual who processes personal data "in the course of a purely personal or household activity" (household exemption). It is a legal challenge to clarify the question regarding the enforcement of privacy rights of users that act as data controllers when publishing personal data about others on the Internet and whether this activity is subject to data protection law and what the consequences are [Kor09]. This question is discussed within the Working Paper 163, Opinion 5/2009 on online social networking of the Article 29 Data Protection Working Party [Art09b, p. 5 ff.] and states that in most cases, users are considered to be data subjects.

But some activities of a user of a SNS may not be covered by the household exemption and the user might be considered to have taken on some of the responsibilities of a data controller (for example, when the social network is used as a collaboration platform for an association or company) [Art09b, p. 5 ff.]. In these circumstances, the user needs the consent of the persons concerned or a legal basis for the processing of personal data.

The directive furthermore states, that also a high number of contacts in a social network "could be an indication that the household exception does not apply and therefore the user would be considered as a data controller" [Art09b, p. 6]. This shows that the Art. 29 Working Party inclines to make users who uploaded content to a wide audience responsible as data controllers. In conclusion it can be said that users of social network sites or blogs, uploading materials for the dissemination to "an unrestricted number of people" are not covered by the household exemption [Kor09].

Most of the users are not aware of the fact and the related duties. They need to learn and need to get information when acting within the web or especially SNS. Therefore there should be guidelines for the user for acting in conformity with the law as well as from the other perspective on how to contact a data controller and how to exercise the rights under the Directive.

Self-determination and controllability of personal data does also relate to portability of data. This may mean that the data subject can control her personal data in a way that they are portable within the web. For example there could be mechanisms that make personal data in a profile of a social network compatible to another social network. In this case the user could "move" one profile to another social network if she wants. This could either be after quitting the participation in one social network and implementing the profile into a new social network or even having identical profiles in different social networks without creating it completely new within the registration process. Profiles could also be exported to the local system of the user (for example, for archiving). It is an important feature for Privacy Throughout Life to preventing "lock-in" situations, i.e., if the user is factually dependent on the data controller (for example, the social network) and cannot leave even if she does not agree with the privacy policy in place.

> *Control-Req l):* Data controllers should prevent lock-in situations. For example, SNS providers should provide portability for user profiles.

Joint responsibility of personal data raises the risk that the data subject loses control about her personal data. Therefore it is necessary to clearly define responsibilities in case of joint responsibility.

> *Control-Req m):* Data controllers, and in SNS also peers, should clearly define responsibilities in case of joint responsibility of data as well as the rules for jointly or separately using the joint data (for example, in a (privacy) policy or another binding contract).

Data subject rights. The fact that data subjects do leave digital footprints in the web, also means, that the data subject has certain rights on the data she left within the digital footprint, such as the general rights stipulated in the Data Protection Directive [Eur95, Art. 12ff].

Apart from the data subjects rights, the discussion point here is to what extent the disclosure of data by the data subject implies consent for processing of personal data and whether these public data are available for all purposes as fruits of the public domain. This question can be answered with the provisions of the Data Protection Directive where

Art. 6 states, that personal data must be collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes [Eur95, Art. 6]. The data subject's consent can be implied for the purposes that are visible for the data subject when giving the consent. If personal data is further processed, even as fruits of the domain, there may be a change of purpose that requires a new consent of the data subject or any other legal basis. Even publicly available personal information has to be used carefully. If somebody wants to further process these data he needs to assure that this processing is legally allowed (legal basis needed), otherwise the protection of personal data may be undermined.

Most of the data subject's rights are stipulated under the provision of the Data Protection Directive [Eur95], but also the provisions of the ePrivacy Directive [Eur02] may be applicable in cases where electronic communications services are provided. If the ePrivacy Directive is applicable, further rights of the user may be taken into account, for example, Art. 6 (4) whereas the service provider must inform the user of the types of traffic data which are processed and of the duration of such processing for the purposes determined. Furthermore it has to be kept in mind that the definition "user" in the ePrivacy Directive means any natural person using a publicly available electronic communications service, without necessarily having subscribed to this service [Eur02, Art. 2a]. In most of the cases the data subject and the user coincide and both provisions are applicable.

It would be useful in many situations to have a "right to start over". The data subject undergoes different phases of life that comprise different kind of personal data. Mostly, the risk of processing personal data is minimised beforehand by legal regulations or technical measures. But in some situations it is not possible to cover all risks and subsequently user control of the processing of his personal data is not possible. This might be the case if wrong personal data are used without one's fault (for example, governmental or third party fault). This incorrect information may cause various problems for the data subject in the daily life and the data subject may furthermore seek for rehabilitation (for example, if wrong personal data appear in web search engines) when the incorrect information puts the person in the pillory.

Therefore the data subject needs to have the "right to start over" in certain situations and should furthermore have the possibility of rehabilitation.

In a lifelong context it becomes even more difficult to keep track of what others (might) know about an individual. The data subject has no possibility to make it comprehensible what others may know or what kind of personal data someone else has. The problem that rises in this context is the handling and organisation of the massive amount of data in case of supporting the user in keeping track. There is also no approach on how to make such data accessible or what kind of interfaces could be used. The data subject does not have any solution how to deal with outdated file formats or hardware components, regular backups and technical updates of his track application. As all these questions still remain unresolved, it is also unclear if additional legal regulations are needed especially with regard to prevent the abuse of data track information by spying or by others unlawfully asking for access to this information. For the data subject it is necessary to also be informed who has obtained certain personal data, for example, data from the electronic personal eID or the electronic health insurance card and where processing of personal data occurs. Therefore the data controller should offer an infor-

mation system for the user to create transparency on where and what kind of data were processed by whom. Data controllers should make this information available for the user and readable with common systems (for example, a list containing who has asked for what kind of data of the electronic health insurance card and for what purposes or a token that may read and tell the user who has the bank account number or the date of birth).

> *Control-Req n):* Data controllers should provide the appropriate information to the data subject to create transparency of what kind of privacy-relevant data is processed by whom. Further they should support data subjects in exercising their rights, e.g., by lowering the threshold to get access to personal data via online solutions.

The data subject does not have full overview over personal data that exist and that are part of the digital footprint. In this context it is also important that with regard to the digital footprint data as total may be personal, even if some data within the digital footprint may not be personal. This means, that data that are not qualified as personal data as such are part of the footprint and in the context of the footprint they are qualified as personal data because they refer to an indirectly identifiable individual. Even if there is no identification with the name or address or social security number, the footprint is identification based on the possibility to single out an individual. For this reason it is even more important for the data subject to be informed about the digital footprint.

3.1.4 User-Controlled Identity Management

In general, the data subject shall have full controllability of all data and purposes within the full lifecycle. A subject's data should be protected for life. This means that each data item should be traced during its life-cycle. When creating data items or accounts and starting up processes, the potential impact on other partial identities should be measured and presented to the data subject for evaluation (evaluation can be partially automated or automatically documented). Moreover, controllability assumes that the information, presented to the data subject is understandable. Finally, deletion should be anticipated, and the desired degree of deletion determined: complete deletion assumes that copies held by data controllers are also deleted, and that secondary usage might not be allowed for such data.

The essence of PrimeLife's approach to identity management builds around the postulation of data subject centricity. The aim is to put the data subject of (new) information technologies (in an online world), e-government services, and offline services facilitating processing of personal data in control of the data processing occurring. The approach of user-controlled identity management as well as of exercising informational self-determination presupposes that the acting data subject fully comprehends the effect of the data processing in question. As described above, transparency is an essential prerequisite for exercising the right of informational self-determination. In order to understand the information given, make a decision as to allow or prohibit the intended data processing and act accordingly and voice this decision, a certain degree of sanity as well

as mental maturity is required. With regards to fundamental rights it is possible to distinguish between a "legal capacity to bear a fundamental right" (Grundrechtsfähigkeit) and "the ability to exercise a fundamental right on one's own" (Grundrechtsmündigkeit).

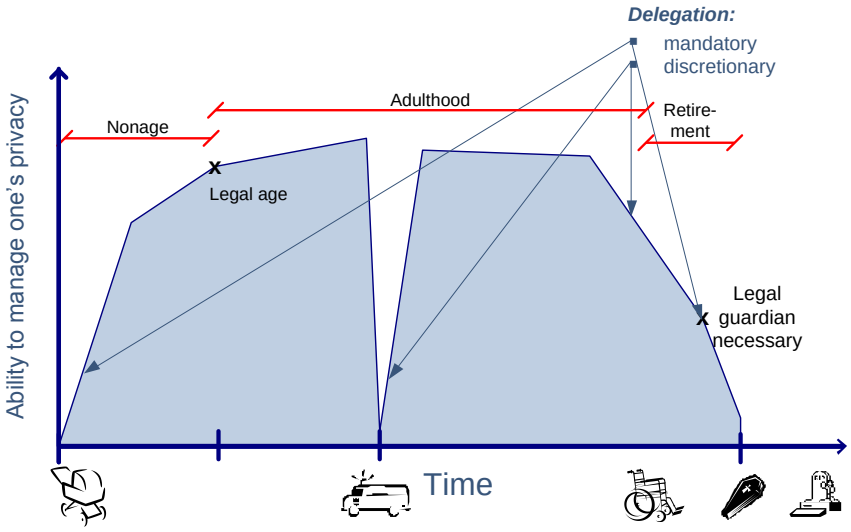

Figure 2: Exemplary stages of life (based on [CHP+09]).

Every natural person bears the fundamental right of informational self-determination. However, every natural person during his or her lifetime passes through (a) stage(s) during which he does not have the ability to understand the consequences of data processing conducted by data controllers, or he is not capable to exercise self-determination via the provided means, for example, due to usability problems. In general, one's life can be classified in three large stages of childhood, adulthood and old age, as shown in Figure 2 [CHP+09].

3.1.5 Delegation in Identity Management

In the context of identity management throughout life, one focus lies on investigating the necessity of delegation for people who are not able to manage their needs of privacy for a limited time or forever. This section describes the general and existing concepts of delegation and derives requirements for delegation in identity management:

Delegate-Req a): Data controllers, data processors, and system developers should foresee that data subjects can delegate their identity management to proxies.

Delegate-Req b): Data controllers, data processors, and system developers should enable delegation of identity management limited to specific proxies and specific scopes (such as purposes, applications, data controllers, time etc.).

Delegate-Req c): Data controllers, data processors, and system developers should enable revocation of delegation of identity management under defined conditions.

Delegate-Req d): Data controllers, data processors, and system developers should provide mechanisms for a data subject to get an overview of decisions by her proxy regarding processing of personal data.

Delegate-Req e): Data controllers, data processors, and system developers should provide concepts and mechanisms for identity management after one's death.

In the following it will be differentiated between the common terms of delegation, as defined above and as legally defined and the definition which is more under the civil law aspect, namely delegation based on explicit decision/will of the data subject. As in the common term of delegation this is mostly a stage of life in which the data subject is not capable to exercise her rights, delegation based on explicit decision/will of the data subject refers to stages of life in which the data subject explicitly wants to transfer full or partial legal authority of representation to another individual.

Delegation based on legal provisions. As mentioned above, the data subject needs to be represented by another natural person who exercises the right on behalf of the data subject concerned during certain phases of life. This may start when a child is born and it may continue in case of adults that may have temporary or permanent needs to get support, and it may finally end with the dead of the data subject's last will. Each stage has significant question on how to handle identity management and in particular personal data and therefore has different requirements. It is quite clear, that a baby is physically less able than a 10 year old to interact with technical devices. But at least small children are not able to decide on their own which data are created and processed and how their private sphere can be controlled. Fundamental law does not explicitly allow for representation by others. Fundamental rights are by nature non-transferable, personal rights.

Relating to stages of life and the handling of one's private sphere furthermore raises the question if the above mentioned legal regulations are sufficient for the data subject to also exercise the right of informational self-determination. In some legal delegations, for example, in case of a contract, the proxy has to process personal data of the individual represented (e.g., in Germany §28 I BDSG Bundesdatenschutzgesetz). This case is legally correct under the civil law, but also has consequences for the fundamental right of informational self-determination what leads to the question if fundamental rights are transferable to a proxy in general. This may be a problem with personal fundamental rights. The right of informational self-determination may be defined as such in some cases and raises the question if personal fundamental rights are transferable in general. If the answer is positive, it may be necessary to find new legal regulations or instruments that stipulate the intercourse with such cases. This also leads to the fact that providers have to supply appropriate technical infrastructures on a legal basis.

Therefore delegation in privacy issues should be recognised by law as far as legally possible, for example, requiring actions in person only where private law acknowledges similar requirements (like the requirement that a will cannot be made by a proxy could correspond with a regulation that privacy rights for the post-mortal period require a specific mandate). It must be compulsory for data controllers to accept declarations of the proxy.

The above mentioned definition of delegation can be derived from this analysis. Delegation furthermore means the transfer of power of legal representation of one natural person to another natural person. This transfer of power can either result from provisions which lay down legal prerequisites or from the concerned natural person's decision. The delegation of exercising fundamental rights on behalf of the bearer of the fundamental right is as such not known in current legal frameworks as fundamental rights are non-transferable personal rights. Legal representation does however impact fundamental rights as a secondary effect.

Mapping delegation technologically a number of requirements can be derived:

Usually delegation is expressed by issuance of a credential ("mandate", attribute certificate) to the proxy. Among the important procedures to be specified are: issuance of the mandate to the proxy, invocation of actions under the name of the principal with the mandate, verification of the mandate, revocation of the mandate from the proxy and expression of acceptance of the mandate by the proxy [PSDCP08].

> *Delegate-Req f):* Data controllers, data processors, and system developers should provide mechanisms for issuance of the mandate of the proxy, invocation of actions under the name of the principal with the mandate, verification of the mandate, revocation of the mandate from the proxy and expression of acceptance of the mandate by the proxy.

Delegation has to be enabled without transferring the original credentials (such as tokens or certificates) of the principal to prevent identity theft. Possible implementations include derived credentials for proxies or that the proxy uses own credentials to get access and then indicates that he acts on behalf of the principal.

> *Delegate-Req g):* Data controllers, data processors, and system developers should support derived credentials for proxies or that enable the proxy to use own credentials to get access and to act on behalf of the principal.

Actions taken by a proxy must be traceable for the principal, for example, by writing into the data track of the principal or granting the principal a right to access the relevant information in the proxy's data track. Also the data track of the proxy should indicate the fact of having acted as proxy and which data was released. However, in case of minors, as principals the logging requirements must not overstrain the capabilities of average parents.

Delegate-Req h): Data controllers, data processors, and system developers should provide mechanisms that allow the principal to trace actions taken by the proxy.

The principal must be able to declare preferences and conditions to the power of proxy, for example, to partially or absolutely restrict certain disclosures, to stipulate preferences or by giving a general guideline for data usage inform of preferences but allowing an exception for a certain transaction he is interested in regardless of the data required.

Delegate-Req i): Data controllers, data processors, and system developers should provide mechanisms for the principle to declare preferences and conditions to the power of the proxy.

The proxy's own desires for maintaining her privacy have to be considered in addition to the privacy requirements of the principal. Data minimising solutions, for example, by anonymous authorisations, can help preserving the privacy spheres of both parties involved.

Delegate-Req j): Data controllers, data processors, and system developers should provide mechanisms to maintain the proxy's private sphere.

The following subsections exemplarily analyse some stages of life in order to show how the management of one's private sphere with respect to handling her privacy may work.

Fruit of the womb. Privacy throughout life comprises a very early stage of life, the prenatal phase of an individual. Even in this stage of life there might be the need to protect personal data, for example, considering the privacy implications of prenatal DNA tests. In many EU Member States there are discussions about the issue of genetic analysis and the threat of using genetic data poses for individual's right of informational self-determination as well as potential discrimination. Regulations regarding requirements for genetic analysis and the use of genetic data could be a solution.

Children and teenagers. Growing autonomy is an important issue in protection of children's rights, in any area of law. The complexity of situations involving minors is based on the fact that children, despite having full rights, need a representative to exercise these rights – including their privacy rights.

Data protection for children starts within the first days after birth and the processing and storage of birth data or medicine data within the hospital. The protection of personal data of children resides more or less in the responsibility of parents or legal guardians. But when a child grows up, other responsible persons for data processing in different areas of life may become involved, such as teachers, doctors or supervisors [HPS08].

The rights of the child, and the exercise of those rights – including that of data protection, should be expressed in a way which recognises both of these aspects of the situation [Art08]. Until a certain age children have no way to monitor data processing, simply because they are too young to be involved in certain activities. If their parents

decide, for example, to put the child's pictures on their profile in a social network, it is the parents who make the decision about the processing of their children's data and give the consent to do so on behalf of the child. Normally, putting pictures of another person in a social network profile requires consent of that person, the data subject. In the situation described here, the parents are entitled to express the consent in the name of the child. Such situation may put the parents in the double role – of data controllers while publishing their child's personal information open on the web, and, at the same time, of consent issuers as the child's representatives. This double role may easily lead to conflicts. Parents must take great care not to cross the line of the child's best interest when processing the child's data.

It is necessary for the parents or other representatives to listen carefully to the interests of the child at least beginning from a certain age and consider those interests when making a privacy-relevant decision as that decision is binding for the child [Art08]. When the child reaches legal age, it may want to change the recent decision of the parents. Therefore the child needs to know what decisions about processing of personal data were made by the representatives. Afterwards the child needs to give her explicit consent for the processing of personal data. This may be implemented in certain operations in a way that the operator is reminded that the person is over 18 and now the explicit consent is needed. This is relevant in many circumstances, for example, medical matters, recreational activities of the child, school matters, or agreements made by the parents before the child's majority.

As children and teenagers are in the process of developing physically and mentally, the rights of the child and the exercise of those rights – including the rights of data protection – should be accomplished in a way which recognises these aspects of the situation. Especially the adaptation of the degree of maturity of children and teenagers is a central aspect that has to be taken into account by their parents. Children gradually become capable of contributing to decisions made about them. It is natural that the level of comprehension is not the same in case of a 7-year-old child and a 15-year-old teenager.[6] This, in particular has to be recognised by the children's representatives. Therefore the children should be consulted more regularly by adults, teachers or caretakers about the exercise of their rights, including those related to data protection.

The children's representatives should also think about a way to document privacy-relevant decisions so that the children or young adults can later easily understand what personal data have been disclosed to whom and under which conditions. They also may then choose to actively approach certain data controllers to give or revoke consent concerning data processing or to request access, rectification or erasure of their personal data.

Adults lacking privacy management capabilities. For adults that may have temporary or permanent needs to get support or that others act on behalf concerning decisions on their private sphere, we distinguish between delegation for legally relevant actions and non-legally relevant actions. All legally relevant actions regarding process-

[6]The level of comprehension is defined in different ways. For instance the US-American Children's Online Privacy Protection Act (COPPA, Title XII – Children's online privacy protection, SEC. 1302) defines a child as an individual under the age of 13.

ing of personal data are based on national legal regulations such as delegation or legal guardianship.

In case of non-legally relevant actions, such as help with a social network or the Internet in general the person concerned can freely decide what to do. The principal could choose a proxy (for example, a caretaker) to act in the name of the person on the basis of a contract to manage the private sphere. Then the person concerned should clearly define her expectations and needs regarding the representation and the power of disposal.

Deceased people. In situations where a person has deceased, the instrument of law of succession applies. The European Data Protection Directive 95/46/EC assigns the right of privacy and data protection to "natural persons" (Article 1). Deceased persons are no longer regarded as data subjects. Protection against an unregulated processing of data concerning deceased individuals in some European legal frameworks[7] is provided by means of a "post-mortal personality right". In some situations, the instrument offered by the law of succession might not be sufficient – further regulations are needed.

For instance, some users of social networks want their profile to exist even after death or at least would like to be informed how the provider handles the personal data and the profile after death. Here the action of providers of social networks is required to find mechanisms and concepts for the handling of profiles after death of the user. Various mechanisms are thinkable, for example, the user could determine how her profile should be handled after death within the registration process (deletion, blocking, proxy to contact, etc.). Therefore, SNS providers need to define clear measures and concepts to determine the handling of profiles after one's death. In some situations even the autonomous action of the SNS provider might be essential for the protection of users. For example if a SNS user dies and the press accesses the SNS site to copy pictures, contacts, etc. of the dead user, the provider has to balance the protection of the users rights and her competence to, for example, block the profile without the consent of the legal assignee (because this has to happen very quickly).

Meanwhile new services appear on the market, which offer to send out secure messages to friends after the death of the user. Their goal is to give people a safe way to share account passwords, wills and other information. When users book the service against payment of a fee, they get options for when to send messages or to delete some messages permanently after their death. It is problematic if authentication credentials of the user have to be transferred to the service which opens the way to misuse because it is not distinguishable for others whether the user or the service acts.

Delegate-Req k): Data controllers should define how to deal with the data subject's data after her death. In particular, SNS providers should define and provide mechanisms for the user to determine the handling of profiles after her death.

[7] Such as Germany: so-called "Mephisto decision" of the German Constitutional Court; BVerfGE 30, 173.

Delegation based on explicit decision/will of the data subject. The civil law knows the instrument of legal representation also for cases where the concerned individual is fully in possession of his/her mental capabilities and decides on his own to transfer the exertion of rights to another person (for example, Articles 172 et seq. German Civil Code[8]). Various reasons exist why a data subject may wish to transfer the full or partial legal authority of representation to another individual. For example a person may simply be unavailable for a longer period of time with no access to information technology which would allow transmitting and enforcing remote decisions (for example, during a scientific or recreational journey to a secluded region). Or a data subject may feel that certain services which are handled online are better to be understood by friends or even a professional data custodian. Actions of and decisions by the authorised representative may have consequences also for the fundamental rights of the principal who at first glance delegated, for example, only the authority to the agent to close one contract on his behalf. Delivering the contractual duties however will possibly also require the processing of personal data. The legal authority to represent a principal in closing a contract does include the implied authority to initiate the data processing steps necessary to fulfil the primary goal. The instrument of legal representation based on the data subjects declared intention may also have effect after the data subject's death. The data subject may during his/her lifetime lay down a last will which binds the heirs. This last will may also comprise decisions regarding how to treat documents or electronic files containing personal data.

The Art. 29 WP defined in its Option 2/2009 [Art09a] principles regarding exercising the right of children. These principles may also be helpful for determining principles on delegation in general, because proxies may have the problem that delegation in privacy relevant situations might be interpreted in different ways. This means that one may have different needs on good practice of handling privacy.

3.1.6 Practicability of Mechanisms

Usability can be defined as one of the general principles. Mechanisms integrated in identity management systems can only help individuals if they are easy to apply. Interfaces have to be well comprehensible for data subjects. If personal data are stored in many different contexts, provided they are all well protected in functional differentiation, how is control and oversight maintained? Provided having an identity management system of full support of partial identities, it still seems to be very hard to differentiate the different partial identities and to avoid linking. A challenge will be to simplify the view for the data subject on her partial identities, the performed transactions and (potential) linkage of disclosed data without oversimplifying. This could mean to hazard consequences of wrong privacy-relevant assumptions.

In general, there might be certain conditions for mechanisms. Mechanisms need to be practical, viable, functional, helpful and useful for individuals to prevent further risks because of mistakes in the data processing and for the exercising of one's rights.

The following requirements primarily address data controllers as those are the responsible parties concerning data processing. However, the requirements should be seen

[8] English translation of Bürgerliches Gesetzbuch: (German Civil Code): http://bundesrecht.juris.de/englisch_bgb/englisch_bgb.html.

as guidance for developers of mechanisms even if they are not involved in the daily operating of the ICT systems. Note that it is not required that all mechanisms work in the online world, but there may be workflows for identity management which do not use computers at all.

Mech-Req a): Data controllers, data processors, and system developers should develop, provide and use the appropriate IdM mechanisms for all parties involved in privacy-relevant data processing.

Mech-Req b): Data controllers (in particular application providers of IdM systems) should ensure that for all parties involved in privacy-relevant data processing the appropriate IdM mechanisms are accessible.

Mech-Req c): Data controllers (in particular application providers of IdM systems) should ensure that for all parties involved in privacy-relevant data processing the appropriate IdM mechanisms are effective, i.e., having the desired impact within a reasonable time frame with a reasonable effort.

Mech-Req d): Data controllers (in particular application providers of IdM systems) should ensure that for all parties involved in privacy-relevant data processing the appropriate IdM mechanisms are transparent concerning their potential impacts, limitations and side-effects.

Mech-Req e): Data controllers (in particular application providers of IdM systems) should ensure that for all parties involved in privacy-relevant data processing the appropriate IdM mechanisms are transparent concerning their effective impacts, limitations and side-effects.

Mech-Req f): Data controllers (in particular application providers of IdM systems) should ensure that for all parties involved in privacy-relevant data processing the appropriate IdM mechanisms are usable for the specific user group (for example, by well comprehensible user interfaces, limitation in complexity etc.).

Mech-Req g): Data controllers (in particular application providers of IdM systems) should ensure that for all parties involved in privacy-relevant data processing, the devices bearing the IdM mechanisms have an appropriate security level (including hardware, operating system, software etc.).

Mech-Req h): Data controllers (in particular application providers of IdM systems) should ensure that for all parties involved in privacy-relevant data processing, the management of (potentially) personal data has an appropriate security level concerning long-term storage, backup and recovery.

Mech-Req i): Data controllers (in particular application providers of IdM systems) should ensure that for all parties involved in privacy-relevant data processing, the appropriate IdM mechanisms will also work in a – due to long-term effects – potentially changed environment and prohibit lock-in risks (for example, by migration strategies, ensuring long-term portability where needed etc.).

Mech-Req j): Data controllers (in particular application providers of IdM systems) should ensure that for all parties involved in privacy-relevant data processing, there are fallback solutions in case the appropriate IdM mechanisms fail or are not accessible.

3.1.7 Dealing With Changes – Change Management

When enabling identity management throughout life, one has to take into account how to deal with changes in society, law and technologies. This not only relates to the data subject, but also affects data controllers and processors. Data controllers, for example, have to ensure legal compliance over time as well as the state of the art in ICT security by implementing data protection management systems. The question here is how appropriate reaction to social changes may be enabled with regard to technical and legal aspects. Changes have to be recognised and collected before new technologies may be developed or new regulations may be stipulated to ensure quality assurance.

ChangeMng-Req a): Data controllers, data processors, and system developers should monitor changes in society, law and technologies and react appropriately (for example, by evaluating chances and risks, adapting current processes, regulation or standards to the changed conditions).

The Directive lists six potential legal bases for data processing [Eur95, Art. 7]. Mostly the processing of data bases on a contract of user and controller or the consent of the user. This raises the question what happens if the legal basis changes. As the data controller is liable for the legal compliance of the processing of personal data, he has to install data protection management processes (in addition to security management processes) to monitor and react to possible changes [Mei09]. Among others, the controller may have to inform the data subject about the change of contract and has to ask for a new consent to the changed contract.

3.1.8 Conclusion

This Section shows that in selected situations analysed above, concrete technical and legal requirements can be derived. These requirements impact different and sometimes also multiple actors that may implement the requirements within their systems. In many situations there is still much to be done to improve identity management throughout one's whole life; the implementation of the requirements and, related to this, the improvement of identity management is mostly aligned to the interests of the actors involved. Therefore, law makers and technical developers need to cooperate to not only adjust, but also to enforce the above mentioned general principles.

3.2 Tools and mechanisms

This Section focuses on technological issues with respect to Privacy throughout life. As the value of privacy-enhancing technologies becomes more and more accepted (cf. [Eur07]) and user-controlled identity management systems have been proposed to solve the challenges of maintaining one's privacy [LSH08], applying these concepts to Privacy Throughout Life seems to be promising. However, this is easier said than done as we will argue with preliminary remarks. Requirements for user-controlled identity management systems to maintain lifelong privacy are sketched. Basic building blocks for that exist in principle, which lists important technical primitives and tools. This section also analyses long-term issues of these primitives and tools which would have to be considered when employing them in user-controlled identity management systems. This analysis yields further requirements to technical concepts and solutions.

3.2.1 Preliminary remarks from a technological perspective

Over the past five decades, tremendous advances in information and communication technologies have substantially facilitated to collect, store, combine, and process information. Whenever such information relates to human beings, data processing affects the informational privacy of the respective persons. So advances in technology are the root cause of many privacy problems our information society is facing today.

However, interference with people's privacy does not necessarily stand in a direct relationship to the level of technological development, but it more depends on the actual design of systems, protocols, and infrastructures. This latitude has fueled the idea to cure the problems created by technology with more technology, as first mentioned by Paul Baran [Bar65] in the 1960s. Nowadays, the term privacy-enhancing technologies (PETs) refers to technical building blocks for systems that are designed to avoid privacy problems without constraining the system's functionality unnecessarily [GWB97]. So ideally, privacy-enhancing technologies should help people to extract all benefits from technological advances without experiencing the negative side-effects on their privacy and individual freedom.

This sounds too much like a panacea, so it is appropriate to ask how much we can expect from privacy-enhancing technologies in general; and in particular when we consider privacy throughout life.

Computing technology became available to governments in the 1940s, to large enterprises in the 1960s, and to end users in the 1980s. Since then, the field has changed very rapidly: typical depreciation periods range from three years for hardware to about five years for software. Maintenance of legacy systems turned out to be very cumbersome and costly. So effectively, even experts lack solid experience with large systems running for more than two decades. Moreover, the existing experience with long-running systems is almost exclusively drawn from closed architectures, physically shielded from the outside world and administered by professionals. So it cannot be generalised to security technology for open distributed networks, which are exposed to a much wider range of threats. Here, the typical latency between the release of the latest patch and the next successful break is usually counted in days (sometimes hours). Hence, looking ahead, it is unrealistic to expect that consumer security technology will reliably protect people's privacy in common computer-mediated social interactions over a lifetime. This statement will probably remain valid in the foreseeable future, unless major scientific discoveries substantially change our conception of computation and information.

Moreover, even if perfectly secure privacy-enhancing technologies existed, its security would be bounded by the weakest link: the user. It is unrealistic to assume that average citizens are always capable to use privacy-enhancing technologies in their own interest without making serious (and irrevocable) mistakes [AG05].

Taking both limiting factors together, the prospects for a long-term privacy-friendly information society by technology are very dim. Carrying the matter to the extremes, two pure strategies for a society to deal with this situation come to mind:

1. Turn back the clock, abolish freely programmable computing devices (or substantially limit access to them) and provide only ICT systems with limited functionality where compliance with data protection law is enforced, or

2. give up claims for (long-term) privacy in large parts of social interactions.

The first option is so unrealistic that nobody discusses it seriously – and there would be several drawbacks, too. The second option does not square well with normative notions of privacy as a fundamental right and it may impose social costs in the long run. Although difficult to quantify, these costs include lost freedom, fewer innovation through conformity, reduced competition, and possibly resource misallocation due to overt discrimination. So none of the pure strategies seems to be a passable way forward. Instead, one might ask if there exists a "mixed strategy" that reaches a better social outcome than either of the pure options; and if so, what can be the role of privacy-enhancing technologies.

Nevertheless (even imperfect) privacy-enhancing technologies are relevant, though the focus on core technologies might differ somewhat.

3.2.2 User-Controlled Identity Management Systems for Privacy Throughout Life

Looking at identity management (IdM) and in particular at user-controlled identity management systems, [HPS08] have elaborated important requirements taking into account Privacy Throughout Life. These requirements which address developers of IdM systems as well as application providers , are summarised in the following:

IdM-Req a): Developers of IdM systems and application providers should provide mechanisms to represent data such as attributes and attribute values in the user's identity management system.

IdM-Req b): Developers of IdM systems and application providers should provide mechanisms to establish, evolve, and use partial identities from personal data such as attributes and attribute values.

IdM-Req c): Developers of IdM systems and application providers should support third-party certification of attribute values of partial identities in the user's identity management system.

IdM-Req d): Developers of IdM systems and application providers should support (privacy-enhancing) reputation systems in the user's identity management system.

IdM-Req e): Developers of IdM systems and application providers should support authentication of actions w.r.t. partial identities.

IdM-Req f): Developers of IdM systems and application providers should support the user in deciding which attributes and attribute values are revealed to whom.

IdM-Req g): Developers of IdM systems and application providers should support users to store and make easily accessible the history which attributes and attribute values have been communicated to whom in which context.

IdM-Req h): Developers of IdM systems and application providers should support delegation concerning all or specifically selected actions, contexts, and/or partial identities.

IdM-Req i): Developers of IdM systems and application providers should support migration to other technologies, i.e., migration to other user devices and other communication infrastructure as well as use for new applications.

IdM-Req j): Developers of IdM systems and application providers should maintain usability so that users can avoid errors as well as perceive their own digital life as continuous.

These requirements which are in line with the requirements in the chapters before, but are limited to user-controlled identity management systems only, name already a few technological concepts. Many of these concepts are already part of the PRIME's blueprint of a user-controlled identity management system [LeSH07][9], some have been added to reflect interactions among peers (for example, the reputation system), long-term aspects (for example, the necessity to make migration possible and refrain from lock-in effects), or stages of life (for example, the support of delegation). The following section deals with technical primitives which are basic building blocks for user-controlled identity management systems.

3.2.3 Important technical primitives and tools

We differentiate technical primitives[10] and tools according to the following criteria:

1. The parties involved: Who is involved and what are their functionalities/abilities?
2. The purpose: What requirements does the primitive achieve for what information?
3. The attacker model: Against whom should the information be protected and who needs to be trusted?
4. The long-term problems: Which problems arise when the system is in use for a lifetime of an individual or even beyond?

It is especially important to not only consider these technical primitives and tools as technologies which already solve many challenges concerning Privacy Throughout Life, but to apply the long-term perspective to them as well, in particular to show potential risks and conditions for their usage. This leads to further requirements when employing these technical primitives and tools which we point out in the following section.

Encryption schemes. Encryption schemes protect the confidentiality of the content of a text (but they do not protect communication-conjunctures if this text is sent, for instance who sends it from, where, when, to whom). There are two types of encryption schemes, the symmetric and the asymmetric scheme. Both types have three phases (key generation and possibly distribution, encryption, decryption): One symmetric secret key for encryption is created and distributed at least to the encryptor and to a possible decryptor in the first phase of symmetric encryption schemes. In the second phase, she encrypts the content to protect with this key. And in the third phase the decryptor (who might be the same person as the encryptor) decrypts the encrypted content. In asymmetric encryption schemes in the first phase a pair of public and private key is created by the decryptor who distributes the public key to possible encryptors who want

[9] The FP6 project "PRIME – Privacy and Identity Management for Europe" is the predecessor project of PrimeLife. See also http://www.prime-project.eu/.

[10] The descriptions of primitives and tools are partly based on descriptions we already elaborated for the FP6 Network of Excellence "FIDIS – Future of Identity in the Information Society" (http://www.fidis.net/).
Please note that we do not intend to write lecture notes on cryptography here so the summaries are pretty short just to introduce the schemes that are used later on for the tools. Detailed information of these concepts can be found in numerous books about cryptography.

to send messages to her. In the second phase, an encryptor encrypts the content to protect with this public key. Finally, the decryptor who holds the private key decrypts the encrypted content in the third phase.

1. The parties involved: There are an encryptor and possible decryptors of the message.
2. The purpose: Thereby both types of encryption schemes reach the following two properties:
   a) Confidentiality of the content;
   b) Unlinkability of encrypted and decrypted content for unauthorized entities.
3. The attacker model: No attacker can break the confidentiality of the content (as long as the encryption scheme is not broken).
4. The long-term problems: If cryptographic assumptions are made, the key length has to be chosen carefully to provide protection for a long time.

There exist numerous implementations for encryption schemes; the most widely known symmetric one might be the onetime-pad and the most popular assymetric one RSA [RSA78]. To ensure unlinkability for unauthorised entities the encrypted data needs to be salted with a sufficient random value and with a magnitude of 256 bit or more.

Secret sharing. (k, n)-Secret sharing was invented independently in [Sha79] and [Bla79] and means protocols for splitting secrets into n parts, called shares, which are distributed amongst a set of several participants. The secret can only be reconstructed when a subset A of the participants with k (with $k \leq n$) of the shares combine their shares; individual shares do not reveal any information on the secret. Generalised Secret Sharing as proposed in [BL90] overcomes the limit of (k out of n) but allows generic monotonic access structures to a secret. Monotonic here means that whenever a set A is sufficient to reconstruct a secret that also a set A' containing all members of A can reconstruct the secret.

1. The parties involved:
   - a so-called dealer, i.e., the initial owner of the secret[11],
   - shareholders, i.e., the participants who get shares, and
   - a reconstructor, i.e., the party that reconstructs the secret.
2. The purpose: For the secret that has to be protected secret sharing reaches a balance between the following two properties:
   - Availability: even if some shares are lost, the secret is not.
   - Confidentiality: an adversary who gains access to only a few shares has no advantage in guessing the secret.

[11] In general, the secret can also be generated in a distributed way so that no single entity ever knows the secret.

3. The attacker model:
    - Regarding availability reconstruction works correctly, if dealer, reconstructor and the necessary shareholders participating in the reconstruction are honest and the communication between them has not been tampered with.
    - Regarding confidentiality any subset of shareholders not containing all of the ones in A gain no information about the secret as long as the dealer and the reconstructor are honest and the communication between them is confidential.
4. The long-term problems: Availablity of the share holders might decrease if no recursive structure is applied. In the case of a recursive structure the (non-existing) relation between the original owner of the secret and the share holders in a recursive structure might be critical for confidentiality and availabity.

Attribute-based encryption. In an attribute-based encryption system as introduced by Sahai and Waters [SW05], a user's private keys and encrypted contents are labeled with sets of descriptive attributes. Every particular private key can decrypt a particular encrypted content only if there is a match between the attributes of the encrypted content and the user's private key.

1. The parties involved: An encryptor and possible decryptors of the content.
2. The purpose: Attribute-based encryption schemes reach confidentiality of content against everyone who does not fulfil the attributes the content is labelled with.
3. The attacker model: No attacker can break the confidentiality of the content (as long as the encryption scheme is not broken).
4. The long-term problems: If cryptographic assumptions are made, the key length has to be chosen carefully to provide protection for a long time.

Commitments. A commitment scheme allows one party to commit to a secret (fix it so that it cannot be changed) without telling another party about it for a certain time. After telling the other party the secret this party is able to verify that this was the secret the first one committed to. Commitments were first invented as unnamed primitives in other protocols, for example, zero-knowledge proof systems, and only later recognised as something that deserves a name because it occurs so often. The first systematic treatment can be found in [BCC88].

1. The parties involved: A so-called committer and a recipient of the committed secret.
2. The purpose: Commitment schemes have a first phase after which the committer is committed to a secret, but the recipient cannot see it yet, and a second phase for opening and verifying the commitment. Thereby it reaches the following two properties:
    - Committing property: The committer cannot change the secret after the first phase.

- Confidentiality property: The recipient does not learn anything about the secret during the first phase.

3. The attacker model:
   - Regarding the committing property even a dishonest committer cannot open one commitment in two different ways.
   - Regarding the confidentiality property the first phase does not give the recipient any information about the secret.
4. The long-term problems: If cryptographic assumptions are made, the key length has to be chosen carefully to provide protection for a long time.

Zero-knowledge proofs. Zero-knowledge proofs were first presented in 1985 by Shafi Goldwasser, et al. [GMR85]. With a zero-knowledge proof one party is able to prove to another party that a statement she made is true, without revealing anything other than the truth of the statement.

1. The parties involved: A so-called prover of the statement made and a verifier of the proof.
2. The purpose: A zero-knowledge proof should fulfil the following properties:
   - Completeness: The prover can convince the verifier of correct statements.
   - Soundness: Not even a dishonest prover can convince an honest verifier of wrong statements.
   - Zero-knowledge: None who interacts with the prover gets any new knowledge about her statement except she explicitly reveals information on it.
3. The attacker model:
   - If the prover is dishonest and her statement is false she cannot convince the honest verifier that it is true.
   - Even if the verifier is dishonest he cannot learn anything other than the truth of the statement proved by the zero-knowledge proof.
4. The long-term problems: If cryptographic assumptions are made, the key length has to be chosen carefully to provide protection for a long time.

Blind signatures. Blind signatures allow one party (recipient) to have a message signed by a second party (signer), whereas the second party is neither able to link the signature with the protocol session during which the signature is created nor with the identity of the original party holding the message and the corresponding signature.

1. The parties involved: A signer and a recipient and possible verifiers of the signature.

2. The purpose: In contrast to traditional signature schemes blind signature schemes have five instead of three phases: In the first phase (the key generation and distribution) the signer creates public and private key for a digital signature scheme and distributes the public key. In the second phase (the blinding) in contrast to traditional signature schemes the text to be signed, is generated by the recipient of the signature (not by the signer) who blinds it (usually by encryption) and sends it to the signer. In the third phase the signer validates that the received input corresponds to the expected content. Even thought the signer is not able to read the content from the blinded input directly, he can verify that the content matches the expectation by utilizing "cut-and-choose" or "zero-knowledge" protocols. In the fourth phase (the signing) the signer signs the blinded text and sends it to the recipient. In the fifth phase (the un-blinding) only the recipient knows how to un-blind the original text and is able to transform the signature to the blinded text to a signature to the un-blinded text. In the sixth phase (the verification) everyone who knows the signer's public key can verify if the signature fits to the text. Thereby blind signatures reach the following two properties:

   - Unlinkability of blinded and un-blinded text as well as unlinkability of the signatures to them.
   - Integrity of blinded text and un-blinded text by the signatures on them.

3. The attacker model:

   - None except the recipient knows the linkability of text and blinded text resp. blinded signature and signature.
   - No attacker can break the integrity of the text resp. blinded text as long as the signature scheme is not broken.

4. The long-term problems: As cryptographic assumptions are made, the key length has to be chosen carefully to provide protection for a long time. The necessary public-key infrastructure has to be sustained for a long time.

Pseudonymous convertible credentials. A credential system is a system in which data subjects can obtain credentials from organisations and demonstrate possession of these credentials. Credentials usually are assigned to pseudonyms. With convertible credentials the data subjects are able to transform a credential issued to one of her pseudonyms to another one of her pseudonyms. This concept was introduced in [Cha85].

1. The parties involved: Users and organisations.

2. The purpose: In an anonymous credential system organisations know the users only by pseudonyms. An organisation can issue a credential to a pseudonym, whose holder can convert this credential to another pseudonym of hers. Then she can prove possession of this converted credential to another organisation and the following properties hold:

   - Integrity of the converted credential.

- Unlinkability of credential and converted credential and thereby unlinkability of the pseudonyms they are used with.

3. The attacker model:
   - Regarding integrity it should be impossible for a user and other organisations to forge a credential of another organisation for the user, even with an adaptive attack on the respective organisation.
   - Regarding unlinkability an organisation cannot find out if two pseudonyms belong to the same user as long as the user does not tell it.
4. The long-term problems: As cryptographic assumptions are made, the key length has to be chosen carefully to provide protection for a long time. The necessary public-key infrastructure has to be sustained for a long time.

In [CaLy01] such a credential system is called anonymous. This term might be misleading because the system does not reach anonymity directly, but only pseudonymity by the use of pseudonyms and unlinkability. This might result in anonymity, but does not necessarily do so if person pseudonyms are used.

Pseudonyms. Pseudonyms act as identifiers of subjects or sets of subjects. Whereas anonymity on the one hand and unambiguous identifiability on the other are extreme cases with respect to linkability to subjects, pseudonymity comprises the entire field between and including these extremes [PH10].

1. The parties involved: The holder of the pseudonym and the parties she uses her pseudonym with.
2. The purpose: Important properties of pseudonyms can include [CK01]:
   - Proof of holdership: Digital pseudonyms could be realised as a public key to test digital signatures where the holder of the pseudonym can prove holdership by forming a digital signature which is created using the corresponding private key.
   - Linkability due to the use of a pseudonym in different contexts.
   - Convertability, i.e., transferability of attributes of one pseudonym to another: The user can obtain a convertible credential (see above) from one organisation using one of her pseudonyms, but can demonstrate possession of the credential to another organisation without revealing her first pseudonym.
   - Authorisations can be realised by credentials or attribute certificates bound to digital pseudonyms, but also in case of digital vouchers transferable to other people by blind signatures (see above) as well.
3. The attacker model:
   - The users can determine the linkability of her pseudonyms hersely.
   - Attacker model of convertible credentials applies to convertability.

- No attacker can break the holdership of a pseudonym and the correctness of authorisations as long as the signature scheme used is not broken.

4. The long-term problems: Pseudonyms used for a long time allow long-term profiles.

Steganography. Steganography is the old art and the young science of hiding secret information in larger, harmless looking files, the so-called cover data. The main difference to cryptography is: If good cryptography is used, the attacker notices that she cannot understand the cryptotext and will hence presume that the communication is confidential. But if good steganography is used, the attacker will think that the cover data is a plausible message which he completely understands. She does not notice any confidential communication. The young science of steganography uses computers to embed secret data for example in digitalised pictures, video signals or sound signals. According to Kerkhoffs' principle, the security of steganographic systems must not depend on the secrecy of the steganographic algorithm but on a key used to parameterise the embedding. Symmetric keys distributed before exchanging secret messages can be used to control the embedding process itself. To increase security, cryptographic systems can be used to encrypt messages before embedding [ZFK+98].

1. The parties involved: A sender who embeds the secret message into the cover data, and a recipient who extracts the message.

2. The purpose: The purpose of steganographic systems is to hide not only the secret message, but also even its existence. This is helpful if confidential communication is suspicious, unwanted or even illegal.

3. The attacker model: An attacker must not be able to decide with probability better than random guessing whether suspected data contains steganographically embedded messages or not.

4. The long-term problems: If cryptographic assumptions are made, the key length has to be chosen carefully to provide protection for a long time. The attacker must not get a better model of the cover data as the sender.

Secure logging. Secure logging of a system's events is needed to find evidences for privacy abuses caused in the processing of personal data.

1. The parties involved: Data controller and possibly also the data subject.

2. The purpose: The data controller is interested in logging all actions which support the impression that she behaves according to the privacy policies, but she is definitely not interested in logging any action which could be taken as evidence for abusing the personal data of the data subject. The data subject is interested in accurate and complete logs, but particularly in those entries which could be an evidence for the abuse of her personal data.

3. The attacker model: Forward integrity for log entries assures that previous entries cannot be altered, even if the system is compromised. Additionally the deletion of

log entries should corrupt all subsequent log entries to reach completeness of log data.

4. Long-term logs pose an increasing risk on the respective users' privacy.

[BY97] introduce the application of message authentication codes (MACs) in order to reach forward integrity. They divide the timeline into several epochs and use different keys for the message authentication in each epoch with the authentication keys destroyed at the end of each epoch. The authentication key for each epoch is derived from the key of the previous epoch. The derivation of new keys has to be done by a non-reversible mapping such that the attacker is not able to reverse this step and obtain a key of a previous epoch. However an auditor is able to check the authenticity of all log entries by reproducing the MAC keys from the the first authentication key. This protocol makes changes in the log entries apparent to the auditor, but does not help out when the attacker deletes log entries within an epoch.

To reach completeness of log entries either sequence numbers as proposed by [BY97] or hash chains as proposed by [SK99] can help. In a hash chain for log entries, the hash of the current log entry does not only depend on the content of the log entry, but also on the hash of the previous log. Thus, the hash of a log entry would only be reproducible as long as the hash value of the previous log entry exists (and is valid).

The data controller has a clear advantage over the data subject by means of deciding on which actions to log. [AB07] suggest using trusted computing in order to assure that continuous and non-selective logging of all events is performed by the data processor. Once personal data have to leave the trusted environment, secure logging can only provide privacy evidences for the leakage, but not for any further processing.

In general, logs of data processing can be understood as metadata of the processed data. Thus, the less logs exists the less mechanisms are necessary to protect the (meta) data against unauthorised access (or leakage) and the easier it is to reason for privacy properties of a protocol or system. Logging and the need for audits is in fact, whenever involving personal data, causing new privacy issues that need to be addressed in a careful manner in order to let the protocol or system benefit and not suffer from the logging.

Linking the technical primitives to the requirements. The high-level requirements from Section 3.1 have not been designed to be implemented solely or predominantly by technological means. So it is no surprise that Table 2 shows that the technical primitives do not cover all areas the requirements address. Most of the primitives stem from the PET core theme of data minimisation. However, they do not stop when implementing data minimising functionality, but also address fair use issues and could be part of user-controlled identity management functionality. In addition, they may be employed in privacy-enhancing feedback mechanisms which support change management on a societal basis.

Although the requirements for user-controlled identity management systems (cf. Section 3.1.4) address more directly technical concepts, the attempt to assign the primitives to those requirements reveals the different layers of the approaches. It is true that all technical primitives can play a role in user-controlled identity management systems, and this is easy to understand for basic and widely distributed modules such as encryption tools

	Transparency	Data minimisation	Fair use	User-controlled IdM	Practicability of mechanisms	Change management
Encryption		x	x	x		
Secret sharing		x	x	x		
Attribute-based encryption	x	x	x			
Commitments		x	x	x		x
Zero-knowledge proofs		x	x	x		
Blind signatures		x	x	x		x
Credentials		x	x	x		x
Pseudonyms		x	x	x		x
Steganography		x		x		
Secure logging	x		x	x		

Table 2: Linking technical primitives to high-level requirements

or for the core technologies for user-controlled identity management such as pseudonyms or pseudonymous convertible credentials. Also secure logging on trusted devices is clearly important for managing reliably one's partial identities. Other primitives help to implement specific functionality, such as secret sharing supports delegation or deciding on post-mortem or emergency access rights to one's partial identities. The following section discusses briefly additional requirements when combining technical primitives to tools or modules employed in user-controlled identity management systems.

3.2.4 Challenges when employing technical primitives for Privacy Throughout Life

The analysis from the previous section shows that all mentioned technical primitives are vulnerable in the long term, and it is hard to imagine primitives in the realm of technology without that vulnerability. Technical progress over time (also including attack technologies) not only requires algorithms and architectures to be upgraded, but also impose a burden on current designs. While concepts as logging and digital signatures can be extended by binding them to the timeframe in which they were generated, communicated data can be recorded for future attacks such as advanced data mining or breaking of encryption algorithms. This means that any data, communicated over a network that allows for interception, can be exploited at a certain point in time. For cryptographic techniques, the horizon of what we can predict is rather limited, because of possible theoretical breakthroughs. Experts are suggesting algorithm/key size combinations for long-term protection (approx. 30 years); higher security levels are targeting "the foreseeable future" [BCC+08]. It should also be noted that the (possible) development of sufficiently large quantum computers will reshape the entire cryptographic field. Another remark is that for signatures, a refresh mechanism can be used to "update" digital signatures (mostly done by time-stamping). Such a mechanism does not exist for encrypted data.

Robustness and resilience of cryptography is therefore discussed in the ICT security community [BMV06]. It is not sufficient to choose a long key size for cryptographic algorithms if attackers may find other possibilities to break the codes. Moreover, if a cryptographic module which is part of a larger ICT system becomes insecure or vulnerable against attacks, this incident has to be dealt with. The design of the ICT system could integrate diverse cryptographic modules with different algorithms where it is unlikely that both fail at the same time. Then the system could switch to the other module. Of course this switch may also be a vulnerability of the ICT system – imagine an attacker switching the system to the weak protection level before attacking it. Currently there are very few products which contain already multiple diverse cryptographic modules to enhance its robustness.

In any case the technical primitives have to be built together and to be orchestrated by ICT systems such as a user-controlled identity management system. Not only cryptographic challenges will occur, but also possible interlinkages and dependencies between different primitives or tools may be problematic. Even updating certain modules may affect the interoperability of the components. Also migration to other systems should be supported which is not trivial either.

In addition to the technical difficulties in the interplay of components, the real-life

settings may pose severe challenges. One challenge comprises data variety. Assuming that the user-controlled identity management system does not cover all potentially privacy-relevant areas of life, the amount of data available to possible attackers who aim at identifying its pseudonymous user cannot be controlled.

Another real-life challenge is based on today's (and future) hard- and software. Due to the rapid progress of technical development it is hard to guarantee a sufficient level of protection against attacks. Even if future technology will cover actual forms of attack, new forms will be created. The approach of solving such issues by equipping data subjects with trusted devices controllable for themselves only might be in contradiction to some states policies having discussed the necessity of backdoors for law enforcement or secret services for several decades.

The three examples of real-life settings above illustrate the difficulties while trying to take all possible challenges for privacy throughout life into account.

3.3 Conclusion

This chapter shows that there are various high-level requirements on what should happen with personal data and what should not happen with personal data. PrimeLife's work on Privacy Throughout Life has shown that there are no ready-made concepts that convincingly solve the challenges of maintaining privacy throughout one's life. In fact, it seems to be a small area in academic discussions only. The shortsightedness of developers of ICT systems as well as application providers can be explained by the short- or medium-term requests on the market. In addition, stability and security of today's technological solutions have to be improved for current purposes before long-term concepts will be tackled. Privacy-enhancing technologies and most of the sketched technical primitives can be rarely found in present ICT products. Thus, developers, providers, and users have to gain more experiences in the potential, usage, and also shortcomings of PETs. However, policy makers should be aware of the challenges of Privacy Throughout Life and plan ahead (see Section 5.2).

It has to be pointed out, that these selected scenarios are not exclusive and represent only some possible aspects where further improvement is necessary within identity management throughout life.

Chapter 4

Demonstrator to Show the Interplay Between Scenarios

The prototype to be developed shall be able to demonstrate the main concepts and features of privacy and identity management throughout an individual's life. This means that the prototype has to measure up to the theoretical findings in this field. During the previous two years, a couple of articles and reports were published [HPS08, CHP+09, BRS+09, SHP+09, PBP10], which sketched the problem space of lifetime aspects when managing privacy and identity. In order to serve as reasonable demonstrator of those issues, the prototype is required to exhibit the main characteristics of it. These are described below.

Accordingly, the foremost features the prototype will have to cope with are the different stages of an individual's life, his full lifespan as well as the different areas of life a user is acting. In this regard, the authors identify mechanisms relevant for

- *user-controlled privacy-enhancing identity management*: handling and management of partial identities, data minimisation, enforceable rules and policies for data processing, and transparency,

in general, and for

- the *areas of an individual's life*: history logging, awareness support, trust and reputation, knowledge and ability to perform typical workflows, interfaces to legacy and emerging systems;
- the *stages of the individual's life*: handling of all delegation-related processes and data, support for different types of delegation as well as
- the *individual's full lifespan*: long-term storage and handling of identity-related data (availability), assurance of long-term robustness of cryptographic protection (confidentiality and integrity),

in particular.

[CHP+09] bring dynamics into play, which have a direct implication on "lifelong protection of individuals concerning their privacy in an ICT-based society". They distinguish between

- dynamics in the surroundings of an individual and
- dynamics in an individual's ability and willingness to manage his private sphere on his own.

In the following, when talking about the first category of dynamics we refer to *external dynamics*, whereas *internal dynamics* are referred to when we discuss dynamics of the second category. Those two categories of dynamics have been looked at both from regulatory and technological perspectives in [CHP+09].

The main problem regulatory institutions currently have to cope with is that law can only react on detected consequences of advances in the processing and analysis of personal data. This means that, as privacy-related issues of new technology are not always possible to foresee, threats to privacy will happen before law is set into position to regulate the issues. Nevertheless, the European privacy legislation (Directives 1995/46/EC and 2002/58/EC) state three important legal principles, which data processing has to comply with and which imply data processing over longer periods and spanning different areas of life:

1. the *proportionality* principle – data processing is timely limited to "no longer than is necessary for the purposes for which the data were collected or for which they are further processed" (Art. 6 (1e), Directive 1995/46/EC),
2. the *data minimisation* principle – "minimising the processing of personal data and of using anonymous or pseudonymous data where possible" (Directive 2002/58/EC), and
3. the *purpose binding* principle – personal data must be "collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes" (Art. 6 (1b), Directive 1995/46/EC).

According to [CHP+09], technological challenges within the field of lifelong management of privacy and identity mainly comprise robust cryptographic security able to cover an individual's full lifespan as well as assuring potentially required migration of the privacy and identity management system to new hard- and software. Further, the different dimensions of sensitivity of attributes have to be regarded when handling personal data. One of the most challenging concepts that becomes eminent when dealing with different stages of life is delegation. For this, different mechanisms need to be covered, e.g., granting and revoking delegations, accountability as well as transparency of the delegation to the "outside".

4.1 Prototype Ideas and Specifics of Them

Within the PrimeLife project, several internal deliverables (those are called Heartbeats) were created. They analyzed the different aspects of the given research area and determined requirements to be fulfilled when managing privacy and identity management

throughout an individual's whole life. In this course, PrimeLife Heartbeat H1.3.4 specifically collected prototype ideas aiming to implement these aspects [BRS+09].

Due to the huge problem space and given the available resources, it is impossible to design and develop a system that is able to cover the whole problem space of privacy-enhancing identity management throughout individuals' whole lifespan. So, H1.3.4 tried to reduce the complexity of this vast area by structuring it along three general scenarios – Digital Footprint, Growing and Shrinking Autonomy, and Digital Estate – each representing a small part of the problem space. Nevertheless, the scenarios were still quite large and required a lot of interaction between various components and infrastructure to cover them in their entirety. So within each scenario, the scope was further narrowed down to 2 – 3 very concrete prototype ideas. The three scenarios and associated prototype ideas from H1.3.4 are depicted in Table 3.

Scenario	*Prototype ideas*
Digital Footprint	Show my Digital Footprint Remove my Digital Footprint Central Data Handling Repository
Growing and Shrinking Autonomy	Passing SNS[1] Sub-Profiles onto Kids Assisted Living Lifelong DataTrack and Delegation
Digital Estate	Secret Sharing File-System Post-mortem Notary Service

Table 3: Scenarios and prototype ideas (cf. PrimeLife Heartbeat [BRS+09])

Digital Footprint. Three prototype ideas belong to the scenario Digital Footprint. One is to give users a tool to gauge the size and shape of their digital footprint (*Show my Digital Footprint*) and to visualize it by different categories. Related to this, *Remove my Digital Footprint* demonstrates an interface to automatically generate rectification or deletion requests for parts of the data in their footprint. Obviously, such reactive mechanisms suffer from weak enforceability, so one step forward could be proactive control of the data handling policies, to which data controllers should obey. The prototype idea *Central Data Handling Repository* helps users to keep an overview of the policies they agreed upon with various services, and assists them in dealing with changes to these policies.

Growing and Shrinking Autonomy. The scenario Growing and Shrinking Autonomy covers all aspects where users (temporarily) lack the ability to actively manage their own privacy. In this context, *Passing SNS Sub-Profiles* onto Kids illustrates how parents can control personal information concerning their children in social software and, when the children have grown up, pass it on to them. Similarly, *Assisted Living* shows how

visions of computer-assisted care can be realised while retaining as much self-control and privacy as possible for elderly people (or patients). On a more general level, *Lifelong DataTrack and Delegation* demonstrates how various forms of delegation to proxies can be handled in a secure and privacy-respecting manner. The prototype idea focuses particularly on data traces created through delegation. It suggests solutions to the delicate question under which party's control such traces should reside after the delegation relation comes to an end.

Digital Estate. The third scenario, Digital Estate, serves as basis for two prototype ideas that show options how to deal with personal information after the death of the respective data subject. *Secret Sharing File System* describes an implementation of Shamir's secret sharing scheme for key recovery. It allows to distribute parts of a master secret (e.g., a password or private key) to a circle of trusted persons, possibly facilitated by making use of social network relations established over social networking services. In contrast to the grassroots approach, *Post-mortem Notary Service* comes up with a demonstrator for a service that might take over the role of notaries in storing, interpreting, and enforcing a person's testament with respect to his or her digital estate.

4.2 Approaching the Prototype

Within PrimeLife Heartbeat H1.3.4, a discussion on the suitability of the proposed prototypes for demonstrating the throughout-life problem space has been started. For this, the Heartbeat authors refer to related concepts, which were introduced in [DoW08] and which the prototype should address. Above all, the prototype to be built should contain the mechanisms showing long-term aspects of identity formation and evolution. History of (partial) identity handling (Lifelong DataTrack) and Delegation are essential concepts for covering the different stages of an individual's life as well as his full lifespan. Related to this are Policies for long-term access and control as well as the consideration of Long-term aspects of sensitivity of identity attributes. Support of an individual's awareness regarding the processing of his personal data and the related policies – especially with regard to the different areas of his life (context awareness) – should also be reflected by the prototype. Further, the prototype is required to offer the possibility to concurrently deal with the dynamics in both the individual's ability or willingness of managing his private sphere on his own (internal dynamics) and his outside world (external dynamics).

Table 4 summarises how the indicated concepts and proposed prototype ideas fit to each other according to [BRS+09]. Those considerations limit the number of possible prototypes as indicated in the following:

- ☞ *Missing important feature(s)*: The first problem that we identified by ranging the prototype ideas in that table is that almost all prototypes (except for "Assisted Living" and "Lifelong DataTrack") would focus mainly on one area of life only. Thus, their applicability in other areas of life is missed though this is one of the major characteristics of the research area and needs to be addressed. The prototypes "Passing SNS Sub-Profiles onto Kids" and "Post-mortem Notary Service" are only singular actions in quite particular stages of life and, thus, do not address dynamics in the surroundings of an individual or the individual itself. Similarly,

the central concept of history of identity formation and evolution is missed within the prototypes "Assisted Living" and "Secret-Sharing File System".

☞ *Existing implementations*: For some of the ideas, either ready or first implementations in form of web-based services already exist. While we should not reinvent the wheel by realising once again the "Show my Digital Footprint" idea, removing of digital footprints is critical with regard to realisation within the frames of PrimeLife project as it lacks consistent communication structures, technical, and legal concepts. Since establishing these concepts requires a lot of effort to be invested in developing mechanisms that are not focal in the sense of the given research area, it was decided not to go for this idea.

Prototypes	**Potential to show dynamics**		**Concepts for privacy throughout life**				
	Internal dynamics	External dynamics	Long-term aspects of using sensitive attributes	Policies for long-term access and control	Delegation of identity and authority	Context awareness	History of identity formation and evolution
Show my digital footprint	x	x	x				x
Remove my digital footprint	x	x	x				x
Central Data Handling Repository		x	x	x			
Passing SNS Sub-Profiles onto Kids			x	x	x		x
Lifelong DataTrack and Delegation	x	x	x		x		x
Assisted Living	x		x	x	x	x	
Secret Sharing File System		x	x	x	x		
Post-mortem Notary Service			x	x	x		

Table 4: Prototype ideas and concepts (based on [BRS+09])

☞ *Limitation to parts of problem space*: The foremost issue with the introduced prototype ideas is that each of them solves only a particular problem, helps to answer

a certain research question, or illustrates how future technology could look like. None of these ideas is actually able to comprehensively cover the concepts of the theoretical framework introduced as key features of lifelong privacy and identity management.

After having drawn these conclusions from the actual prototype ideas, we came to the decision that we need an additional prototype idea trying to cover the majority of concepts, which especially comprises different areas of life, stages of life including the whole lifespan of an individual, and which is able to show dynamics. This led us to the following considerations: In everyday lives, people are interacting with the physical and digital environments. In both of these environments, there are unpredictable events, which we can neither influence nor foresee and which might have an impact on our everyday lives or on lives of our closest relatives. With computerization of society, human beings are not only more and more dependent on the data but they are also becoming data themselves. As far as the influence of the technology on our everyday reality increases, the protection of data and privacy of the corresponding data subject from an increasing number of risk factors is becoming a crucial part of our everyday reality.

Therefore, we decided to design and develop a backup and synchronisation demonstrator specifically respecting and demonstrating privacy and identity management throughout one's whole life. This prototype solves not only the problem of data protection but also the one of protecting privacy of the corresponding individual and, in addition, it respects different areas and stages of the individual's life. Furthermore, our proposed solution deals with the aspect of lifetime and is able to respond to internal as well as to external dynamics.

☞ *Comprehensive approach*: The backup and synchronisation demonstrator not only addresses the key features of lifelong privacy and identity management. It further incorporates a selection of the main aspects addressed by the previous ideas or it can potentially be enhanced in such a way.

Thus, it will take up issues of Lifelong DataTrack and Delegation by allowing for delegation of work items when the primary user is not able to proceed with his work (cf. stages of life). A logged history of data evolution is required when backup data shall be recovered from a specific point in time. The Secret Sharing File System becomes an issue when, e.g., recovery of backup data should only be possible with the help of cooperating participants. The Post-mortem Notary Service may become an important instance if backup data should be possible to recover after the primary user has passed away. Similarly, the idea of Assisted Living could also be linked to the scenario of synchronising and backing up the states of an elderly person with his nursing service and his medical doctor.

4.3 Implementing the Requirements to Come Up with Solutions

Many backup systems and backup strategies, which have been available for many years, are already dealing with the problem of unwanted data loss. However, they are mostly

protecting the raw data only and do not involve the data subject, his specific characteristics, social relations and interactions as a part of their scope. Existing backup systems and backup strategies also do not reflect the process of evolution of the data subject during his lifetime with respect to possible different states he might pass through during his lifetime and which might have an immense influence on his ability to manage his data on his own behalf (e.g., illness, hospitalization, or death). Additionally, existing systems and strategies dealing with the problem of unwanted data loss do not also cope with boundaries among distinct areas of the data subject's social interactions. However, these aspects are nowadays becoming more and more sensible on the level of the data, hand in hand with the massive expansion of the technology.

Therefore, the problem of unwanted data loss from the perspective of lifelong privacy will be analysed. The findings have shown that current solutions do not provide a sufficient level of data protection when it comes to lifelong extent of time and privacy of the data subject holding the data. Based on those findings, it was decided to demonstrate that it is possible to cope with problems amplified by the requirements on lifelong privacy when protecting the data subject against unwanted data loss.

The proposed privacy-enhanced backup and synchronization demonstrator focuses on the following problems closely linked together under the light of lifelong privacy:

1. Protection of the data subject against unwanted data loss during his lifetime by redundancy and physical distribution of the data;

2. Assurance of lifelong confidentiality of the data subject's data stored in a distributed environment;

3. Delegation of access rights to the data subject's backup data allowing other parties to operate with his data if specific conditions are fulfilled;

4. Distribution of the backup data according to different areas of life of the data subject and his different partial identities.

This section puts the high-level requirements elaborated in PrimeLife Heartbeat H1.3.5 ([SHP+09]) into the context of the specific environment of the demonstrator and points out corresponding requirements and implications for the demonstrator adapted to the specific environment. At this point, we want to mention that terms used in the following course are described in detail within the *Glossary* at the end of this document.

4.3.1 Relating the Backup Demonstrator to the High-Level Requirements of Privacy Throughout Life

The aim of this Section is to clarify the relation between the backup domain and the goals of privacy throughout life. The high-level requirements documented in Section 3.1 serve as the basis for the elaboration. These requirements are transformed and adapted into a more specific form here in order to reflect the nature of the application area of the demonstrator.

Transparency

Transparency plays an important role in many areas of our society. In general, transparent behaviour among particular subjects allows those subjects to be informed about activities, actions, results, and other relevant information related to the corresponding subjects behaving transparently. Transparency brings openness, clearness and controllability to relations among interacting subjects. Transparency plays an essential role in those situations, where data processing is being performed such that certain parties are coming in contact with the data related to other parties. It is therefore necessary to consider transparency as one of the key aspects of the WP 1.3 demonstrator.

Transparency in general. In terms of the WP 1.3 demonstrator, the requirement of transparency in general to be realised as follows:

For all parties involved in all backup-related processes[2], it is necessary that they have clarity on the legal, technical and organisational conditions setting the scope of their role with respect to the data or privilege corresponding to their role[3].

The indicated requirement on transparency can be further extended into more specific sub-requirements specified on the level of concrete actors of the privacy-enhanced backup and synchronization demonstrator:

Above all, it is necessary that the primary user becomes familiar with the basics of the distributed backup schema and also with potential risks that are amplified by the nature of the distributed environment. The primary user must be familiar with protection mechanisms (existing on the technical as well as legal level), which protect his data. He must also be able to learn what the services are and guarantees provided by a storage provider and, under which conditions and to what extent, the storage provider provides his services especially when it comes to lifetime aspects and the death of the primary user. It must be clarified that if the primary user takes advantage of services of an external storage provider, he fully relies on the storage space provided by the particular storage provider and his technical equipment, which is not under the physical control of the primary user. The primary user must also understand what are the potential risks and privacy implications when he enables other parties to restore his backup in case that a specific condition is satisfied.

Transparency in the scope of revocability and irrevocability. The high-level requirement w.r.t. transparency regarding revocability and irrevocability can be adapted to the following form fulfilling the scope of the backup environment:

- For all parties involved in the back up, recovery, delegation of access rights, or which provide storage for the backup as well as other third parties involved in the privacy-enhanced backup and synchronization demonstrator schema, it should be clear under which circumstances their actions are revocable/irrevocable and what

[2]These are: the back up, recovery, delegation of access rights, or providing storage for the backup as well as other third parties involved in the privacy-enhanced backup and synchronization demonstrator schema

[3]For example, clarity on regulations such as laws, contracts, or privacy policies, on technologies used, on organisational processes and responsibilities, on data flow, data location, ways of transmission, further data recipients, and on potential risks to privacy

can be the potential impact. In particular, the primary user should be informed what his possibilities to revoke access rights from corresponding delegates are and what impact does a particular deletion action have on the backed up data. The primary user must also be informed about possible ways of deletion of his backup data and about corresponding implications of particular types of deletion. He must further be aware that he can always delete whichever item of his backup data in all existing instances (even older copies of some item) and in all backup locations under control of the primary user. It must be clarified what happens to the backup data on the storage provider's site. It must be explicitly specified if the storage provider utilizes some backup mechanisms and strategies also on the server side and what impact does it have on the primary user's data in case he deletes his backup items or cancels his contract.

- A delegate must be aware of his possibility to refuse access rights delegated by a delegator in any point in time and the delegator must be informed about it as soon as possible.
- Storage providers must be aware of their possibility to cancel the contract with a primary user in case that the primary user violates conditions defined by the storage provider, which are accepted by the backuper during his subscription.

Awareness. With respect to the privacy-enhanced backup and synchronization demonstrator, the high-level requirement on awareness means:

The primary user must be informed that, by delegating access rights to several delegates, his areas of life can be linked together. He must also have clear control about which data can be linked under which conditions. The primary user must be aware of the possibility of linkage of his operations performed on distributed backups. He must be informed which functionality can be provided by using anonymisation service and how it can help him to avoid linkability and other related problems.

Privacy and security breach notification. In terms of the backup and synchronization demonstrator, the requirements regarding privacy and security breach notification can be interpreted in two adapted formulations as follows:

- In case that some attack method on any security mechanism, which is used in the backup and synchronization demonstrator, appears or any security function, which is used in the schema, is considered to be unsecure, the primary user must be informed about existing risks with respect to potential consequences on the privacy and security of his data and provided advice on how to cope with this problem.
- In case of a successful attack on the storage provider, the primary user must be informed about this incident and about possible consequences with respect to his data and how to deal with these consequences.

Data minimisation

The principle of "data minimisation" refers to the requirements that a data controller should limit the collection of personal information to what is directly relevant and neces-

sary to accomplish a specified purpose. The data controller should retain the data only for as long as is necessary to fulfil that purpose. In other words, data controllers should collect only that personal data they really need, and should keep it only for as long as they need it [Eur95].

Data minimisation in general. In the privacy-enhanced backup and synchronization demonstrator, the general requirement of data minimisation pertains to storage providers and delegates and can be phrased as the following:

- The possibility of the storage provider to observe or link actions of the primary user must be avoided or minimised to an acceptable level. This means that a particular storage provider should not implicitly be able to learn the real identity of the primary user. The primary user should not use the same personal identifier for different storage providers because this could lead to linkability of the user's actions especially in case if two or more storage providers are controlled by a single entity.

- Delegates should be able to perform selected operations authorised by the primary user on his backup in case that all access conditions predefined by him are fulfilled. Any delegate should however not be able to link or observe actions of the primary user (backuper, respectively). In particular, the delegate should implicitly not be able to link or observe actions covered by diverse areas of life of primary user or his partial identities. The primary user should be able to keep his areas of life separated and link them only in cases he explicitly wants to link them for one or more selected delegates.

- An attacker must not be able to observe actions of the primary user (backuper, respectively) performed on the backup stored in the distributed environment (on storage provider's equipment). An attacker must not be able to link actions performed by the primary user, which means that he cannot learn that two actions performed on one particular backup were initiated by the same primary user.

Minimal quantity and sensitiveness. Under the scope of the privacy-enhanced backup and synchronization demonstrator, the requirement of minimal quantity and sensitiveness of the data controlled by third parties relates to storage providers and delegates, in the first line. With respect to the backup demonstrator, it is to be interpreted in the following way:

- Storage providers and delegates should have only minimal access to personal and sensitive data of the primary user. As far as the role of storage providers is to provide remote storage space for backups of the primary users only, they have no reason for accessing the content of the primary user's data. Therefore, according to the principle of data minimisation, storage providers should not be allowed to access the backup data of the primary user. They should operate only with such a type of data, which is necessary for providing their services and accounting. This means that confidentiality of the backup data has to be assured as far as the data is stored in a storage space provided by storage provider.

- The access of delegates to the primary user's backup should be minimised to such an extent, which is wanted and expected by the primary user. The primary user should minimise the access of selected delegate(s) to specific backup data by appropriate access conditions. If the primary user decides to minimise access to his backup data by using access conditions, selected delegates are allowed to access the backup data only in case they provide the corresponding credentials issued by credential issuers, which proves that the access conditions are fulfilled.

Moreover, the need to minimise quantity and sensitiveness of the data handled by involved parties is extended in *DataMin-Req c)*. This has influence on the backup demonstrator in the following way:

Security mechanisms, which assure unobservability and unlinkability of the primary user's actions as well as anonymisation and pseudonymisation of the primary user's identity, must be supported by the demonstrator. In case that a storage provider needs to hold some personal data of the primary user for providing his services, this data must be erased as early as possible.

Minimisation of time frame of data storage. Adapted to the privacy-enhanced backup and synchronization demonstrator, requirement *DatMin-Req d)* means that:

- Every delegate who was given access rights determined by access conditions must be allowed to access the data only within the duration of validity of the particular access conditions and for the specific purpose of access rights. This means that as soon as the access conditions are no longer valid, the delegate must not longer be allowed to perform permitted operations on the backup data of the primary user with respect to the particular access conditions.

- From the storage provider's point of view, this high-level requirement obliges the storage provider to minimise the time frame of holding personal data of the primary user with respect to the duration of the contract between the storage provider and the primary user. After this period of time, which must be defined within the terms and conditions and accepted by the primary user, the storage provider must implicitly and immediately erase any identifier or information, which leads to identification of primary user and his backup data as well.

Minimal data disclosure. For the privacy-enhanced backup and synchronization demonstrator, requirement of minimal data disclosure relates to storage providers, delegates and credential issuers. All these actors must minimise the extension of personal data to that level, which is necessary for fulfilling the specific purpose. For example, a primary user (delegator, respectively) should delegate access to only that data which fulfils the purpose of the delegation. This implies that an appropriate access condition should be selected by the primary user (delegator, respectively) according to the purpose of the delegation. In any case, the primary user must be aware of the fact that as soon as the delegates gain access to his data, the primary user (delegator, respectively) has to rely on the trustworthiness of the delegates because in fact he has no longer direct control on what actually happens to his data. Also backup data, which may contain

personal data of the primary user, should not be disclosed to storage providers or any other third party if not authorised by the primary user (delegator, respectively).

Minimal correlation possibilities – limiting linkability. Within the scope of the privacy-enhanced backup and synchronization demonstrator, the requirement to minimise correlation possibilities *DatMin-Req f)* mainly applies to the primary user. The primary user should minimise linkability and linkage of his actions and data by using suitable tools assuring unlinkability, which must be supported by the demonstrator. Above all, linkability should be avoided between different areas of life and different partial identities of the primary user.

For the privacy-enhanced backup and synchronization demonstrator, the high-level requirement *DatMinReq g)* means that the demonstrator must provide mechanisms, which minimise multipurpose or context-spanning use of (potentially) personal data stored in the backup or stored by the storage provider for accounting purposes. In particular it means that storage providers, attackers or other third parties (legally related or not) are not able to use potentially personal data of the primary user for different purposes.

Within the scope of the privacy-enhanced backup and synchronization demonstrator, that *DatMin-Req h)* refers to the requirement that the demonstrator must avoid using unique identifiers in different contexts. A primary user's accounts provided by different storage providers must use diverse identifiers. Any two backups, which were created for different purposes, must also be stored under different accounts using diverse identifiers.

DatMin-Req i) primarily relates to storage providers in terms of privacy-enhanced backup and synchronization demonstrator. Storage providers should support anonymous or pseudonymous authorisation and access control of users where possible. Pseudonymous authorisation and access control should also be supported between primary user (delegator, respectively) and delegates as well as between delegates and corresponding storage provider.

Avoid or limit irrevocable consequences. For the privacy-enhanced backup and synchronization demonstrator, the interpretation of requirement *DatMin-Req j)* leads to the following two requirements adapted for the environment of the demonstrator:

- The primary user (respectively delegator) should always be able to revoke access rights delegated to delegates.
- The primary user should always be able to remove any backup item contained in any backup he created including older backup items of the same primary item.

Fair use – Controllable and controlled data processing

From the point of view of the privacy-enhanced backup and synchronization demonstrator, the first high-level requirement regarding "fair use" *Control-Req a)* concerns storage providers and credential issuers. Processing by the storage provider, as well as by the credential issuer, should be controllable and controlled throughout the full lifecycle and it should be compliant with the relevant legal and social norms.

According to the specific purpose of the demonstrator, it should be assured that the demonstrator provides appropriate mechanisms, which allow the primary user to easily separate his data and create a backup corresponding to the specific purpose respecting potential risk factors during the lifetime of the primary user (*Control-Req b)*).

Storage providers should be specific in the definition of what kind of information is required in order to support accounting and anonymous payment – *Control-Req c).* In case that there are some third parties to whom a storage provider provides information about his clients, this must be explicitly mentioned to the primary user including the purpose for which this information is provided and what the potential consequences are.

Accountability. Regarding requirement *Control-Req d)*, which demands that data controllers should limit the data subject's consent to data processing in time, every access rights of users of the backup and synchronisation prototype should include a reasonable validity period by default.

Control-Req e) additionally requires that data subjects should be able to withdraw their consent without any impacts on their privacy. Adapted to the special environment of the demonstrator, this requirement means that a primary user (respectively delegator) should be able to revoke access right delegated to one or more delegates. On the other hand, the backup and synchronization demonstrator does not allow the primary user to permanently remove any data, which was provided this/those delegate/delegates by delegation of the access right.

According to *Control-Req f)*, a primary user should be able to make the delegate accountable. The primary user should be able to define and assign clear responsibilities, which must be clear to the delegate before he accepts the delegation of access rights.

Additionally with *Control-Req g)*, identity theft needs to be prohibited. Hence, the primary user should:

- prohibit identity theft by not delegating access rights of the backup data to delegates who are not trustworthy regarding the specific purpose of the backup. Sensitive data should be distributed and protected in such a way that no unauthorised person is able to access it.
- avoid identity theft by not providing his real identity to any storage provider if not necessary. This means that there should be a mechanism which allows the user to communicate with storage providers anonymously or at least pseudonymously.

Sensitive Data. Requirement *Control-Req h)* can be adapted into the following form:

- A primary user (delegator respectively) should be extra cautious when delegating access rights to delegates especially in case that the backup for which the access rights are delegated contains sensitive data of the primary user. Delegates should be extra cautious when accessing the backup data delegated by the primary user (delegator, respectively). This holds especially in such cases when a delegate is bound by a legal agreement (for example non-disclosure agreement).
- Storage providers should provide such mechanisms and policies which do not allow any unauthorised third party to access the potentially sensitive data of the primary user (backuper, respectively).

Organisation of Data Processing. A primary user (delegator, respectively) should conceptualise and plan his backup recovery strategy (*Control-Req i)*) resulting in the corresponding access rights and conditions before the access rights are delegated to delegates. During the creation of a new backup, the demonstrator should provide the possibility to define the time period in which backup items are automatically updated on a regular basis. In case that the primary user (backuper, respectively) does not specify a time period for backup updates explicitly, the demonstrator should ask the primary user (backuper, respectively) if a corresponding backup should be updated in case that some primary item was modified.

Control-Req j) requires data controllers to foresee concepts and procedures for erasure of used identifiers. This requirement can be adapted to the formulation that, in case that there are some identifiers created (for instance for the purpose of accountability by storage providers), there should already be concepts and procedures, which assure that those identifiers are erased after the usage period.

Primary users as well as storage providers should be prepared for emergency situations, e.g., security or privacy breaches (*Control-Req k)*). For a storage provider, this could, e.g., be an unrecoverable damage of a storage medium. For the primary user, this could, e.g., be a loss of connection during the backup procedure or appearance of an attack method, which defeats some security mechanism which the demonstrator relies on.

Dealing with Possible Conflicts. According to *Control-Req l)*, lock-in situations should be expected and prevented by the demonstrator. For example in case that a particular storage provider does not provide stable services or is temporarily out of service, the demonstrator should allow the primary user to easily migrate his backups to some other storage provider allowing corresponding delegates to still have access to the backup data. This means that there should be a mechanism, which allows the primary user (backuper, respectively) to move his backup data stored by a particular storage provider to a storage space provided by another storage provider. In a simple setting of the demonstrator, this should be achieved by downloading the backup data directly (if possible) or from corresponding redundant copies of the backup data stored by other storage providers (in case that the affected storage provider is not able to provide access to primary user's data) and subsequently uploading the data to storage provided by the other storage provider. In a more advance setting of the demonstrator, it should be possible to realize direct upload from storage provider(s) to another one without the need to download the data to the primary user's side before uploading it. In any case it must be assured that actions performed during the migration are not linked to each other. The "receiving" storage provider should not be able to learn that the incoming data is coming from the "sending" storage provider and the "sending" storage provider should not be able to learn that the data is sent to the other storage provider as well. The migration mechanism must assure that the information about the current location of delegated backups is updated accordingly after the upload of the backup is finished. In a more advanced setting, the update of the current location of the backup should not require active participation of the corresponding delegates having rights to perform particular actions on the migrated backup.

In order to fulfil *Control-Req m)*, storage providers should clearly define internal

responsibilities and rules for its staff members, especially in case that this particular storage provider relates on equipment, which is physically separated in distributed storage facilities around the world.

Consent and revocation

In general, the users' data should only be accessible to authorised third parties.

In terms of the privacy-enhanced backup and synchronization demonstrator, this requirement applies to primary users (delegators, respectively) and their possibility to delegate access rights of their backup data to other delegates. If it is possible, validity of the access rights delegated by the primary user (delegator, respectively) should be limited in time by default according to the purpose of the backup. For example, access rights delegated to a company, which employs the primary user, should be valid only for the period of time for which this primary user is working for that particular company.

Usability

Data subjects should be made aware of potential risks to privacy and ways to deal with these risks, for example, in privacy policies.

For the primary user of the privacy-enhanced backup and synchronization demonstrator, it means that he should be made aware of potential risks to privacy especially in a case that he stores some data in a distributed environment and in case that he delegates access rights of his backup data allowing other entities to operate with this data.

The evolution of user experience over the full lifespan of the data subject needs to be considered as well. 'Unambiguous human-machine communication' is crucial to keep the elderly and people with low education as long as possible able to act on their own behalf [SHP+09].

This means that the backup and synchronization demonstrator must provide interfaces, which adhere to common usability principles reflecting the specific needs and characteristics of its individual users and user groups.

4.3.2 Socio-Cultural Requirements

This section deals with the socio-cultural requirements for the demonstrator and is based on the socio-cultural requirements derived from the project PRIME – Privacy and Identity Management for Europe[4].

In PrimeLife's predecessor, the PRIME project, a list of socio-cultural requirements was established. These requirements relate to privacy and identity management. A number of them, however, can also be relevant for this demonstrator. First, here the entire list of requirements will be repeated and, then, the relevant ones will be indicated. These requirements should contribute to a refinement of the requirements to the demonstrator.

In the "PRIME Requirements V3" [KDR+08] a distinction was made in three key aspects, the "umbrella terms", which are relevant from a socio-cultural perspective. These umbrella terms are

[4] https://www.prime-project.eu/

- audience segregation,
- control, and
- adoption.

Control is constituted by ten requirements, namely: “comprehension”, “consciousness”, “consent”, “choice”, “confinement”, “consistency”, “context”, “inspection”, “chain control”, and “ex-post user control”.

Adoption is constituted by six requirements: “social settings flexibility”, “minimise skill level”, “accountability”, “trust in transaction partners”, “trust in communication infrastructure”, and “affordability”. Requirements can influence each other and sometimes they overlap.

Audience segregation

Audience segregation is very relevant in the context of the demonstrator. The primary user should be able to have different partial identities to play different roles and portray the self to others in a way she chooses. With regard to the demonstrator, this means that the contents of the backup have to be divided in categories belonging to the different audiences the individual interacts with. Once, the backup system is needed to provide a person (second party) with information or contents belonging to an individual, access to the backup needs to be restricted to the parts that have to be disclosed to the aforementioned second party instead of permitting it to access all the content in the backup and selecting the relevant parts themselves. The issue of audience segregation should be taken care of in the demonstrator by letting the primary user provide others with access rights, either directly or via delegation. These access rights should be connected to specific parts of the content. Besides, the content itself needs to be encrypted, so that visibility for second parties does not directly imply a “real” disclosure of the content. In this way the issue of audience segregation can be solved in the demonstrator.

There is, however, a specific point of attention, which is related to indicating or defining the audience and having control over this. A distinction can be made between the intended audience and the actual audience. The intended audience is the audience which was meant to have access to content in the backup system. So, this is the second party to whom access rights were granted or delegated. The actual audience is the audience that in practice has access to the content. The actual might be different from the intended audience when, e.g. the access rights are distributed to further users or when the originally intended audience has changed in composition. That can, for instance, be the case when a colleague has also become a family member in the meantime and, therewith, has access to work-related as well as family-related documents. Here, the segregated audiences come together and contexts collapse. The consequences of this will depend on the way the colleague deals with this.

Apart from this, it is questionable whether this problem can be solved in the demonstrator anyway. Changes over time in people’s contexts are difficult to grasp in a technical solution. Nevertheless, a proper delegation system may solve the issue. When a third party is responsible for delegating the access rights when necessary, the distribution of access rights can take place at the moment when it is necessary and does not have to

take place (possibly long) beforehand. This implies that the intended audience can be reviewed at the moment of granting the access rights.

Control.

Control means that the user has control over what happens with her personal data. This is relevant not only for the primary user, but also for the delegates as far as their personal data are concerned.

For the primary user[5], this implies the data items for which backup items will be created, the backup items, usage data that is created when using the backup system, information on requests to delegation candidates and delegates and the communication regarding the status of the primary user that may be communicated to delegates. For delegates (and delegate candidates), privacy-relevant data may be usage data as well as requests from the primary users or credential issuers. So not only the data items themselves, but also information on a delegation (candidate) relationship between a primary user and a delegation (candidate) as well as usage data have to be taken into account when dealing with the control principle.

In the light of the demonstrator the control principle means among others that the primary user decides what is stored in the backup, who has access, and can check along the way and afterwards whether things went as intended. The aspect of control is constituted by ten requirements. Each of them will be discussed here briefly.

Comprehension. Comprehension means that the primary user as well as delegates should understand how their personal data are processed by the service provider. For the demonstrator this means that the primary user and the delegates have to understand what happens to their personal data.

Basically, the primary user will often have the role of service provider herself when using the backup tool, so then it is necessary that the primary user understands what exactly happens when she is doing something with her data. Thus, the system and its functionalities itself need to be comprehensible. Next to that, there may occur situations where others process the data, for instance, when the data are stored on another machine than the primary user's own computer. Then, the owner of the other machine can probably be qualified as a service provider. At least there is the service of hosting or providing backup space.

It has to be clear to the primary user whether the tool is running on the primary user's computer after installation or whether the service is or can also be provided by a third party at a distance, who then can be qualified as a service provider.

Consciousness. Consciousness is described as: "the user should be aware of the essential events, processes, stakeholders and attributes of the collection and use of personal data." Consciousness is a necessary condition for the exercise of data subject rights such

[5]In addition to the "primary user" and "delegates", PrimeLife heartbeat H1.3.6 defines the actors "backuper", "restorer" and "deleter". Here we assume that they should be understood as roles taken by the primary user or a delegate rather than being actors themselves. If they were actors, at least their usage data may be privacy-relevant so that all control requirements would have to be considered for them, too.

as consent, right of access and right to object. When a primary user uses the backup system, she has to be aware of the processes involved in the backup and who is involved in the processing of the data. Similarly the delegates have to be aware of privacy-relevant aspects of the backup system – both in the relation with the storage provider and with the primary user.

In the demonstrator, the data subject takes the role of a user and creates the backup. This implies that there is consent and that the data subject is aware of the data being processed. The exact process, however, also needs to be transparent and comprehensible. Only then will the primary user (or a delegate) be in a position to be aware of (possible) second or third parties that may have access to the data or process the data otherwise. Exact technical knowledge may not be necessary, but some knowledge about the exchange of data and access to data is essential.

Consent. As indicated above, the requirement of consent is fulfilled in the demonstrator since the data subject is the primary user of the backup system and initially processes the data herself. Processing by others is based on delegation and access rights, meaning that the primary user also has consented for this (further) processing.

Further the delegates have consented to be delegates, which means that they have to know what is expected from them and how the primary users or others may control their behaviour. Note that the delegation candidates have not given consent to be delegates.

Choice. Choice means that the primary user should have choices regarding all data collection activities concerning her personal data. Taking into account that the idea of the demonstrator is to give the user all control over the processing of her data, this requirement is less relevant. As long as the primary user is not forced by the system to process more data than strictly necessary for the purpose of the backup system, there is no need for an alternative system to provide the primary user with choice concerning the system used for backing up data. The same is valid for delegates.

Confinement. Confinement is in fact a very broad term. It covers the well-known principles of purpose specification and purpose binding – or better even: use limitation. This entails that the primary user should be able to set limits on who may access her personal data and for what purpose. Moreover confinement relates to security safeguards because the user should be able to set limits on who may access personal data. These requirements obviously are also legal rights and duties, but they may be hard to enforce in practice. The PRIME vision is that mechanisms to enforce these legal requirements should be embedded in techniques and applications.

The demonstrator ought to take care of these requirements properly by calling for inter alia minimisation of linkability of personal data (risk of "function creep") and also minimisation of multipurpose or context-spanning use of personal data. Storage providers are thus prevented from using the personal data of the primary user for different purposes (e.g., accounting purposes). Likewise attention needs to be given to security requirements such as planning for emergency situations. Data controllers also have to foresee beforehand procedures for erasure of personal identifiers after their usage period. As such, the

demonstrator will fulfil the following targets of the PRIME requirement of confinement by providing ways to express preferences/policies with respect to:

- the purpose of use of the personal data,
- who may have access to the (personal) data,
- where they may be stored,
- until when they may be stored.

It also should provide the necessary security safeguards to keep personal information within its determined boundaries. This is not only necessary for the data items and backup items, but also for usage data of primary user and delegates and for requests between the primary user, the delegates, and the credential issuers.

Consistency. The requirement of consistency is very much tied up with the so-called digital identity of the primary user. Contextual dependency plays an important role here as well. The context dependence necessitates that the primary user is informed about the use that is to be made of the data she provides via the application, because each application potentially gives rise to new and unknown uses and ensuing identities. These uses and identities still lie in the future at the moment the primary user relinquishes her personal data to the application. In other words, there is a time lag between the principal moment at which the primary user is able to exert control and the moment the digital identity comes into existence. The primary user must be given insight into the future uses of the personal data she provides. This glimpse into the future must be constructed with as much consistency as possible. The principle of consistency is therefore related very much to the requirements of control and transparency.

The demonstrator ought to take much note of this requirement. It should elaborate the transparency requirement from the scope of revocability and irrevocability. It should prescribe that for all parties involved it should be clear at all times what the potential impact can be of decisions and under which circumstances they can be revocable or irrevocable. Data controllers have to inform primary users of the ir-/revocability of their decisions. The primary user must be aware that she can always delete any item of her personal data in all locations and in all existing instances. The primary user should also be able to refuse access rights delegated by a delegator at any point in time.

Context. The requirement of context is linked to the requirement of audience segregation and choice. The caveats noted about the differences between "intended" and "actual" audience can be brought to bear here as well. However, contexts bring into play new elements because context tries to take account of situational factors affecting individual perceptions and desires for privacy. It also takes into consideration the kind of information in question, in terms of its perceived sensitivity. For young people the definition of sensitive information is frequently very different from what is specified in data protection and privacy laws. Not only age differences play a role here but also broader socio-cultural variables. Situational or physical contexts may affect the primary user's privacy preferences to a large degree.

The demonstrator should seek solutions for separating the personal data of the primary user according to her different areas of life because these areas belong to different contexts. All the same the proper delegation of sufficient access rights by secure access control mechanism needs to be implemented. The primary user usually should be in full control of the selection of these access rights; delegators others than the primary users should be the exception (e.g., in emergency cases or on a predefined basis), and even their actions should be comprehensible by the primary user. Of course, the confidentiality of the personal data can be guaranteed by data encryption before sending it to the storage controller. Important measure in this respect will be the utilisation of the anonymisation service which separates the different areas of the primary user's life. Context separation can also be ensured by anonymous credentials unique for every backup.

Inspection. Inspection relates to the control requirement because users must be able to check whether their actions have the desired effects. It is as such a key requirement to establish transparency. Obviously, data subjects have legal rights, such as to be informed about the processing of personal data, the identity of the controller and the purpose of the processing. But these rights need to be implemented to make sense and be effective. As such inspection is a means of ex-post user control.

The demonstrator ought to take care of this requirement by adopting all this type of information in the End User License Agreement which the primary user will have to accept before using the service. In start-up windows, interactive help information during the use of the application, and a wizard, which provides the user helpful information related to the action at hand, much attention can be given to this requirement. Information can also be provided about the risks of linkage and the granting of access rights to delegates. A search functionality can also be included in order to permit the user to visualise which kind of data is stored in which backup.

Note that primary users often wish or even need to check whether the actions of delegates performed on their behalf don't violate what has been agreed upon. And also delegates may desire some inspection possibility as far as their personal data are concerned.

Chain control. The primary user and delegates should be able to inspect data collection and use throughout the service chain. Following from the requirement of inspection, users (i.e. primary users and delegates) should be able not only to inspect the actions of a data collector, but also the actions of the multiple service providers which are present in the interaction chain. As such, chain control is a specification of the inspection requirement.

In general, chain control means among others: On the user's request, the application should provide information for each party involved about:

- When personal data has been disclosed?
- To what parties this data has been provided?
- Under what conditions the data has been provided?
- Who had access to the data?

- For what purpose they had access to the data?
- Why and how data is used?

The application should also allow for informing the user later on about a new link in the chain (for example when a new organisational structure of a business where the user is employed is introduced) and if so, make it possible to change or withdraw consent.

The demonstrator should pay much attention to fulfilling these requirements. This is of special importance in the case of third parties becoming delegates of a particular backup. Every delegate candidate is then informed who initiated a delegation request for what data and under which conditions the delegator will be able to perform what operations on this data. The delegate candidate should be informed that he can refuse (revoke) delegated access rights anytime in the future if he accepts it. Selection of proper access rights should be under full control of the primary user.

Special attention should be paid to the user interface which informs the primary user that the storage provider may cancel the contract if the primary user does not follow the requirements of the contract, e.g., if the primary user refuses the contractually agreed payment. This should only be possible within an accepted and previously communicated legal setting. Perhaps this could result in disclosing the personal information of the primary user to enable the storage provider to take legal action.

Ex-post user control. This requirement is closely linked to the requirement of inspection and, it hardly needs saying, user control. Ex-post user control is control after the fact. Therefore it builds on the legal requirement of the right to rectify, erase or block the data (Article 12 of the Data Protection Directive 95/46/EC). From a social perspective this is an important requirement because personal data are increasingly used as the basis for making decisions concerning an individual. If the data are incorrect, outdated or not relevant, the decision can turn out to be incorrect, and hence, users must have the possibility to correct or object. In addition, ex-post user control is an indicator of how people perceive the service and is therefore related to comprehension and consciousness.

In the context of minimising irrevocable consequences the demonstrator should implement a function which allows for deletion of any backup item and its derivates in whichever backup of the primary user. Solutions should also be provided for backup and removal of a single item taking into account an encryption schema, anonymity, unlinkability and unobservability during transmission. The demonstrator will then provide sufficient opportunity for removal and deletion of data by the primary user. Special attention should be given to also providing a mechanism for rectification.

Adoption

The demonstrator should be designed to maximise adoption by its target audience. Obviously, a tool which is not used is useless. However, not only the use as such is of importance, but also the use by a (large) group. Only when a bigger audience is reached the tool will contribute to improvement of identity management capabilities of individuals and at the same time prevent the risk of a digital divide between users who can manage their identity and personal data and those who cannot.

Social settings flexibility. This issue is probably not or only less relevant here. It deals with perceptions of public and private, but in the prototype all data are primarily considered as private data.

Minimise skill level. This requirement means that the user should be able to use the tool with a minimal amount of training and cognitive load. So, the purpose of the tool has to be clear and the interface has to be user-friendly. This user-friendliness should count for a general public, so not only people with technical skills, but also ordinary users. The demonstrator should strive for a clear and comprehensible user interface.

Accountability. This issue is probably not or only less relevant here, because the tool's primary purpose is not to act anonymously. However, the access to the system based on pseudonyms, either by the primary user or by a delegate, has to be accountable. This is an access control issue which should be taken care of by means of credentials.

Trust in transaction partners. This issue is probably not or only less relevant here – the demonstrator doesn't focus on transactions, but on access to data. Surely trust in delegates is essential for the privacy-enhanced backup and synchronisation demonstrator, but it mainly has to be tackled outside the IT system.

Trust in the communication infrastructure. The user has to be able to assess the trustworthiness of the communication infrastructure. So, how secure is the backup, can others access it or not, how do things work when someone else must access the data? Requirement therefore is that this is dealt with properly by means of an access control system. Access control and authorisation should also be dealt with in relation to security. End-user trust must be provided by usability, also, because usability can contribute to transparency of the system and therewith increase trust. If a system is too incomprehensible (not user-friendly), trust is difficult to establish.

Affordability. The tool should not be too costly for users to obtain and use. If it is a software tool that can be distributed fairly easily, this will be no problem. However, it should also be easy to install in order to prevent the tool from being costly in terms of time spent ("interaction cost").

4.3.3 Privacy-Related Requirements for Delegation

This chapter focuses on privacy-related requirements for delegation in the privacy-enhanced backup and synchronisation demonstrator. After giving definitions from [HRSZ10], the setting of delegation in the WP 1.3 demonstrator is explained. On this basis, several requirements on how to tackle delegation in the demonstrator are identified, in particular on limiting the delegate's access to the necessary extent, on controlling the delegate's actions and in the area of delegation based on legal provisions.

The Setting of Delegation in the Demonstrator

We address privacy aspects of delegation as a means to support individuals in stages of life when they cannot act on their own or are not willing to act on their own regarding some aspects of their privacy although they might be capable to do it.

To clarify our understanding of delegation, we quote the following definitions from [HRSZ10] that are in line with the legal terminology:

Delegation: *Delegation is a process whereby a delegate (also called "proxy", "mandatory" or "agent") is authorized to act on behalf of a person concerned via a mandate of authority (or for short: mandate).*

Mandate of authority: *The mandate of authority usually defines in particular*

1. *the scope of authority for the actions of a delegate on behalf of a person concerned and*
2. *when and under which conditions the delegate gets the power of authority to act on behalf of the person concerned.*

The delegate shall only act on behalf of the person concerned if the delegate has the actual power of authority and if his action lies within the scope of authority. The simple acting of the delegate with the existence of a mandate while not having the power of authority would not be sufficient. The difference between mandate and power of authority becomes clear in the following example: In working life the schedule of responsibilities may determine that person A should take over the work of colleague B if the latter is absent. The issuance of the mandate of authority to A is expressed by the schedule of responsibilities, but the A's actual power of authority only comes into existence if B is absent. Otherwise A must not act on behalf of B.

Delegator: *The mandate of authority is issued by the delegator (also called "mandator"). This may be the person concerned herself, but there are also cases where other entities explicitly decide on the delegation (e.g., in the case of incapacitation of a person the guardianship court rules on delegation) or where the delegation is foreseen in law (e.g., when parents are the default delegates of their young children). The mandate of authority is usually assigned for a specific period of time. Similar to the process of issuing a mandate, changing or revoking the mandate can be done by the delegator, i.e., by the person concerned herself or by other entities. The conditions and processes to issue, change, or revoke a mandate can be defined by the underlying contract or law.*

Note that not always the delegate is aware of the mandate of authority or of the fact that he actually has the power of authority. So the delegator should implement an appropriate way of informing the delegate (and the person concerned if she is not the delegator herself) about the mandate and the power of authority.

Delegation Supervisor: *For supervising purposes of the delegation and related actions by the parties involved, one or more impartial delegation supervisors may be appointed by one or more of the actors. In particular the person concerned may have the need to check whether the delegate really acts as agreed upon.*

The concept of the WP 1.3 demonstrator that is sketched in PrimeLife heartbeat H1.3.6 takes a slightly deviating approach [DB10]:

Similar to the definitions in [HRSZ10], the delegator (which may be the primary user herself) has the privilege to delegate rights to delegates in the H1.3.6 setting. As a mechanism, a delegation request is sent to a delegation candidate who can accept or refuse being a delegate concerning particular rights. From [DB10], it is not fully clear what "particular rights" the demonstrator will support: Presumably it should only mean "reading access" to predefined backup items (under a defined partial identity of the primary user, possibly only the newest last version and possibly for a limited period in time). But it could also mean that the delegate may back up items from the primary user (e.g., if the delegate acts on behalf of the primary user in a transaction, the relevant data could be put into the backup), that the delegate may restore (or copy) the backup on another computer (which may be necessary if the delegate should act on behalf of the user), that the delegate may conclude a new contract with another storage provider, or that the delegate may cancel an existing contract with a storage provider.

Requirements for Delegation

On the basis of the outlined definitions and concepts in the previous section, this section sketches various requirements for the integration of delegation into the concept and possibly implementation of the WP 1.3 demonstrator.

Limiting the Delegate's Access to the Necessary Extent. As already discussed, the delegate's access should be limited to what is necessary. This comprises the limitation to one or few partial identities of the primary user as well as a limitation of the possibility to exercise the access rights to a certain period of time.

The delegator should be supported in limiting the delegate's access rights by the user interface of the backup system. By default, only access to the newest backup as well as to data belonging to one partial identity of the primary user should be offered. The system should inform the delegator that the access could be even further restricted, namely to specific backup items. It should explicitly ask for the period of time that the delegate should be assigned the according access rights. By default it should not be unlimited, but extensions should be possible if necessary (e.g., in the case of a hospital stay of the primary user which turns out to be longer than expected).

The backup system should support the delegator in the following steps:

- How to generate delegation requests to delegate candidates?
- How to deal with their (positive/negative/missing) answers to those requests?
- How to revoke the status of being a delegate?
- How to limit the access rights of the delegate?
- How to communicate possible conditions to being a delegate or conditions to having the actual "power of authority"?
- How to communicate to the delegate that he is assigned the actual "power of authority"?

The question of the actual "power of authority" that a delegate should have if he acts on behalf of the primary user is tackled in [DB10] by issuance of a "credential verifying a certain status of the primary user". For visualising the functionality of the demonstrator, this construction may be sufficient. Still from a data minimising perspective it should be clarified which information is necessary in each case to be transferred and who sees that information (the credential issuer? the delegate? others?). In particular it is very often not necessary, but even privacy-invasive to give information on the medical status of the primary user.[6] In several cases it could be sufficient to communicate "delegate receives the power of authority from `<beginning>` to `<end>`". For delegates it is very relevant to know whether the power of authority is only given for a very short time so that he only has to handle urgent requests and could delay the non-urgent requests, or whether the power of authority should last for a longer time. Even in case the exact timeframe of the power of authority cannot be given in the beginning, this fact as well as a minimal or estimated duration for the power of authority should be communicated to the delegate. This could also mean that the delegate gets multiple messages for prolonging the power of authority or refining the information given in a message before.

Controlling the Delegate's Actions. The primary user should always be in full control over the access possibilities of the delegate and should also be able to reconstruct the actions the delegate has performed on behalf of the primary user. A precondition for this is the information of the primary user of the assignment of delegates (or delegate candidates) and their potential or actual rights concerning the backup. This is important to maintain an overview on the delegation at all times, but it is even more necessary in case the delegator who initiated the delegation is not the same person as the primary user.

It is mandatory that the delegates do not use the same credentials as the primary users to perform their actions because otherwise these actions would not (or not easily) be distinguishable from actions from the primary user. Concerning delegation to organisations where multiple members (e.g., employees) could act as delegate, each individual person should use their own credentials. There should be more than one credential per delegate if they are assigned access rights for different backups (different users or different partial identities).

Actions taken by the delegate concerning the data of the primary user must be traceable by her so that she can check later on any action performed with respect to her data. If she cannot conduct the supervision herself, she may appoint one or more impartial delegation supervisors to look after her interests. For the demonstrator, this means that it should log actions performed by delegates for a predefined time and give access to these logs by the primary user. This has to be known both by the primary user and the delegate. The demonstrator could also foresee the possibility of delegation supervisors that cannot access the data of the primary user, but can access the logfile on the actions of the delegate. Delegation supervisors are delegates, too, but only in the

[6] Note that contrary to the impression imposed by the following quotation, pregnancy as such usually does not have an impact on the ability of the primary user to manage her privacy: "State of life: temporary or permanent state of the data subject's life, which can be certified by a corresponding credential issuer and which might have impact on the ability of the data subject to manage his data (e.g., illness, hospitalization, death, pregnancy, imprisonment and others)." [DB10]

function of supervising other delegate's behaviour. They need own credentials to prove that they have specific rights. In case their actions (i.e., accessing the logfile) should be supervised, too, there would be the need for another logfile. Surely it does not make sense to implement a fully recursive (and thereby infinite) mechanism of logging and supervision. Moreover, there should not arise further risks for the privacy of the primary user – or the privacy of the delegate – by maintaining comprehensive logfiles for a long time. Here a good balance has to be found. There is no need for the demonstrator design to fully resolve this issue, but still it should foresee possibilities to log actions of the delegate and provide access to the corresponding logfile by the primary user.

In case the delegate should be allowed to assign further delegates with the same or derived access rights to the primary user's backup, this has to be communicated to the primary user. This sub-delegation needs to be traceable later on, too. Similarly the delegate has to notify the primary user (as well as the delegator) if his credentials get lost or stolen or if he has the suspicion that somebody misuses his credentials. In case the delegate cannot perform the actions the primary user (or the delegator) expects, e.g., because the delegate's stage of life does not allow it, this fact also has to be communicated to the primary user (and the delegator). It is not necessary that the demonstrator implements such kinds of notification, but the concept should mention the necessity of such functionality.

The primary user should be able to define the scope of authority of the delegate: Under which conditions should the delegate access which data? In addition, there may be specific preferences for post-mortal period that the primary user would like to communicate to her heirs or some delegates. For the purpose of the demonstrator the development of technology-supported mechanism to express conditions or preferences might be too ambitious.

Delegation Based on Legal Provisions. The concept of the demonstrator in PrimeLife heartbeat H1.3.6 ([DB10]) does not elaborate on different causes and procedures how the delegator may assign delegates. However, possible assignments of delegates have been depicted in an HCI prototype within the PrimeLife project [GWK+10]. The presentation of the prototype in that deliverable showed that it is not easy to design a clear and understandable user interface for assigning delegates by the delegator, and the desired functionality has not been fully spelt out, yet. In the following, a few issues related to different causes for delegation are described.

The instrument of legal representation is common in civil law where the powers of the delegate and the legal effectiveness of the delegation are predefined as well as the bounds of delegation. Many of the scenarios depicted in PrimeLife heartbeat H1.3.6 ([DB10]) contain delegation aspects based on the will of the individual. However, the delegation may also be based on legal provisions, e.g., if the delegate is the legal representative of the primary user according to law or a court decision. This is especially relevant if backup items contain not only private diary entries, but something relevant to official transactions (e.g., governmental certificates or insurance documents).

Law defines generic roles (and associated conditions to check that a person is playing that role lawfully) in addition to family relationship in cases where the primary user, over majority age, is unable to manage her own data (e.g., mental disability). Several types of roles and levels of delegation might be defined, with more or less control over the

person's data, in accordance with their role in assisting the primary user in her everyday life.

Delegation by a legal representative based on court order. The first question is how to prove that a person is a legal representative of a primary user. Today, the proof (e.g., a court order) would have to be shown to the storage provider. In the setting of the demonstrator this other party may rather be the credential issuer. However, a prerequisite would be to know that the primary user has used the backup service of specific storage providers.

Current legal concepts usually don't distinguish between different partial identities – instead, the task of a legal representative in the sense of the civil law is to represent a person in all contexts. This would mean to give access to all backup items of the primary user, irrespective of the chosen partial identity. Of course the primary user could plan ahead for possible legal representatives, e.g., by only informing them about specific partial identities or individually encrypting data she does not want to disclose to her legal representative.

Access to backup items from deceased persons. In situations where a person has died, the instrument of law of succession applies [StHR09]. Therefore the legal basis for deceased people is as follows: The European Data Protection Directive 95/46/EC assigns the right of privacy and data protection in Article 1 to "natural persons". Deceased people are no longer regarded as data subjects. Nevertheless, protection against an unregulated processing of data concerning deceased individuals in some European legal frameworks is provided by means of a "post-mortal personality right".[7]

The primary user could determine beforehand how her backup items should be handled after her death (containing order for further storage, deletion blocking or delegation). In this approach, the primary user would give orders similar to a testament. This has the advantage that the way of handling the data would be defined even beyond death. However, in the applicable civil law (e.g., in Germany, section 1922 BGB) this might conflict with the role of the "universal successor" that reserves all rights and all duties from the deceased individual (or simply the inheritors). This typically includes the right to decide what happens to the individual's rights – and her goods. Although this usually does not include the right to determine about the individual's personal data, the access to some backup items may be necessary to know about relevant financial assets or to access digital goods stored in the backup that now belong to the heirs.

Another question would be which influence the death of the primary user could have on the delegation. The mandate – as long as it is not limited to the primary user's lifetime – is still valid when she deceases. The universal successor has to revoke this delegation if he does not want the delegate to have access rights to (certain) backup items of the individual. Therefore it is doubtful, whether the backup system has to inform the delegates about the primary user's death or whether the backup system just would have the obligation to inform the universal successor about existing mandates. Following the principle of data minimisation, the second approach seems to be preferable.

[7] This applies at least for Germany, see the so-called "Mephisto decision" of the German Supreme Court, BVerGE 30, 173.

The universal successor has to legitimate to the backup system, afterwards the backup system has a contract with the universal successor (including all rights and all duties).

It is questionable whether the credentials of the primary user should directly be transferred to the inheritors in case of her death. For example a person from a group of inheritors should be prevented from accessing the data first, copying them and then deleting all the data so that the others don't have access to this information anymore. On the other hand, there may be personal files that are only meant to be read by one specific person within the group of inheritors. In any case inheritors should not work with the original credentials of the primary user, but be issued individual credentials. The possible conflicts (that are usually resolved in different ways according to national law) should not be dealt with in this version of the demonstrator.

Using the backup system for data from minors. Parents could manage a backup of official information concerning their children. Initially, the delegation works from the parents to the child, i.e. the child can have access to the data, but it is entirely managed by the parents until an age defined by law. The topic of delegation for children and teenagers has also been depicted before [StHR09][HRSZ10]. It includes the unusual feature that the delegate does not only act on behalf of the individual, but the delegate also has to decide, which access rights the child or teenager will get and when. In other words, in this scenario the delegate decides about the individual's access rights and not vice versa. At some point, the parents should decide to reverse the delegation: the child then gains control and the parents receive the delegated role. Then, this delegation should not be revocable by the child until the age of majority. In any case, at majority, the grown-up gets full control over the data without the necessity of involving the parents as delegates any more. The young adult also may decide to remove all the data from that storage provider and choose other services.

The scenario is different from the others described before because the law already defines that parents are legal guardians for their children. Insofar they can and should act as delegates for their children. The parents take over the role of a delegate (or rather two delegates) and in addition are in the role of the delegator (again: possibly two delegators) while their children are individual primary users.

Since this would reverse some functionality in the demonstrator, it could be mimicked in a way so that the parents first decide on which data to backup where and give each child the reading access rights as a delegate (although in principle it is a primary user). It should be negotiated whether the access of parents or kids are logged and if so, who gets access to these logfiles. Usually there is no real need for logging the activities (at least not in a trusted relationship), but this may not hold for all occasions.

Privacy-Related Requirements Derived from the Backup and Synchronization Nature of the Demonstrator

Additional privacy-related requirements stemmed from the specific backup and synchronization nature of the demonstrator are introduced in this section. These requirements come up due to two main reasons:

- Backup data of the primary user of the demonstrator is stored primarily in a

distributed environment and in redundant copies in order to deal with the problem of data loss during the lifetime of the primary user.

- Other parties can operate with potentially sensitive content of the backup data, which belongs to the primary user in case predefined specific circumstances are fulfilled.

Location of backup data. The more storage space on different storages located in diverse storage areas is available to the backup and synchronization demonstrator the more robust behaviour of the demonstrator can be achieved. Sufficient amount of distributed storage space enables the demonstrator (or primary user) to create more redundant copies of the primary user's data and also provides more possibilities for separation of different areas of his life and partial identities. However, as far as the amount of data stored in different distributed locations increases, the primary user might lose control over the location of his backup items.

Therefore, the primary user must be able to know what data is stored in which of his backups. He must also have some mechanism, which allows him to visualize which of his existing primary items are backed up so that he has a clear idea which of his data has already been backed up and where. The primary user should also be able to search his backup items based on selection of certain criterions (e.g., based on areas of life, partial identities, date of creation, stage of life when the item was created, name of the item, and others). The primary user should also be able to detect those backup items for which an original primary item no longer exists. When searching for a file or performing any other localization action, any potential attacker must not be able to learn what backup data the primary user or delegate (fulfilling all access conditions) is searching for, in which backup is that particular data stored or even that he is performing a search action. Attacker should also not be able to link a user's identity with the location of the user's backups.

Backup and removal of a single item. The primary user should be able to insert (respectively delete) a single item to (respectively from) an existing backup. The application of this requirement should assure that the primary user can effectively operate with backups stored in the distributed environment. When performing the operations insert or delete on a single item, any potential attacker must not be able to reveal the content of the item inserted or deleted by primary user, what backups are influenced by this action or even that an insert or delete action was performed. Any potential attacker should not be able to reveal any of the backups stored by particular storage providers.

Back-in-time recovery. The primary user (backuper, respectively) should be able to create a backup that allows him to recover previous versions of one or more backup items reflecting the previous state(s) of the corresponding primary item. This means that the primary user (backuper, respectively) should not only be able to recover the last state of the primary item archived by the most recent back up action but also any previous state of that particular primary item created by backing up the item in the past. This functionality should allow the primary user to return back in time and restore previous states of primary items. Previous versions of primary items must be handled in the same

way as ordinary backup items when performing some actions on the backups so that the same level of privacy is assured for all backup items.

When delegating access rights, the primary user (delegator, respectively) should be able to specify whether the delegate (restorer, respectively) should be able to access all versions of a particular backup item (the most recent version and all previous versions created in the past) or only a selected subset of versions of a backup item (e.g., delegate access rights to all versions of the backup item which will be created within next two weeks). The primary user (delegator, respectively) should also be able to delegate access rights to a particular version of a backup item and all of its successors (even those which do not exist in the time of the delegation) or to a particular version of a backup item and its predecessors.

The backup, which stores previous versions of the backup items, might generally reveal more information about the primary user than a simple backup storing the most recent versions only because the former contains the primary user's data spanning a broader time frame. Therefore, it is necessary to consider the fact that, once the backup contains data covering long-term history of the backup items, it becomes more valuable for potential attackers. Thus, advanced security mechanisms assuring privacy of the backup data should be utilized by the demonstrator.

Full deletion. The primary user (deleter, respectively) should also be able to completely delete all backup items created for a particular primary item by using the "full deletion" function. This means that as soon as the "full deletion" is activated on some backup item, all related backup items corresponding to the same (possibly no longer existing) primary item must be deleted. This also includes related backup items stored in different backups of the primary user (stored by different storage providers) as well as older versions of backup items in case of back-in-time recovery function activated. Full deletion of a backup item must assure that for a selected backup item all of its existing forms (including different versions) are erased from all backups and there is no evidence that it was ever stored in any of the user's backups. This function enables the primary user to delete any backup item at any point in time. This function must enhance the primary user's privacy so that, after activation of the full deletion, he is sure that all copies of that particular backup item stored in any of his backups are deleted (even those stored in a backup device of the storage provider in case that the storage provider backs up the client data by default on the server side). The full deletion must be applicable to all backup items, which enables the primary user (deleter, respectively) to delete all of his backup data. The process of full deletion must be compliant with the requirements on privacy, anonymity (and pseudonymity) of the primary user and unlinkability of his actions, partial identities, and areas of life. Full deletion supports the primary user in his possibility to "start over", which is discussed in PrimeLife heartbeat H1.3.5 (see [SHP+09]).

Backup recovery after unrecoverable crash of the user's system. The primary user must be able to recover all of his backup data stored in all of his backups. For example even in a case that the computer of primary user burns and all the local data is permanently lost, he should be able to recover all of his backup items possibly on a

different system.

4.4 Solutions for Relevant Requirements of the Demonstrator

This Section presents solutions for relevant requirements introduced in Sections 4.3.1 and 4.3.3. Furthermore, the line will be drawn between solutions, which will be directly implemented and demonstrated and those which will be solved on a conceptual basis. Further, only those requirements which are directly relevant for the demonstrator from the implementation point of view and which can really be demonstrated as real privacy-related solutions by our demonstrator are considered here.

The solutions to be presented are elaborated on a high-level approach currently not describing concrete detailed technological details of the final demonstrator. More concrete technical and implementation details as well as the according mechanisms will be part of deliverable D1.3.2, which especially focuses on the technical background of the demonstrator.

4.4.1 Solutions for transparency requirements

This section introduces solutions which fulfil requirements on transparency introduced in the Section 3.1.1. The main goal of the "transparency" in general social context is to provide openness, disclosure, awareness or accountability. In this section, solutions for the demonstrator based on transparency requirements are presented.

Solutions for Openness, Transparency, Notice, Awareness, Understanding

The first requirement adapted for our demonstrator requires that it is necessary that the primary user becomes familiar with basic technical background of the distributed backup. Moreover, he should be informed about potential risks of the backup environment and corresponding protection mechanisms as well as conditions, under which storage providers offer their services.

Solutions for the demonstrator. In our demonstrator, this requirement will be solved by stating this information in the End User License Agreement (EULA), which will have to be accepted by the primary user before the installation proceeds. Furthermore, this information will also be presented in the start-up window in an interactive way introducing the above-mentioned information in several well-structured steps. After that, it will also be possible to open an introductory information window by activating the corresponding menu item at any time during the user's work. There will also be a wizard, which will provide the user necessary information related to the action, which the user plans to perform. Also, it will warn the user about possible consequences in case the action of the primary user might have an impact on his privacy.

Integration in the demonstrator. These solutions will be addressed in the form of conceptual specification in the demonstrator.

Solutions for transparency of what is irrevocable and what is revocable

Particular actors playing their specific roles in the demonstrator's environment should have clearness under which circumstances their actions are revocable and under which irrevocable (see Section 3.1.1 for details).

Solutions for the demonstrator. In our demonstrator this requirement will be solved for primary users, delegates and storage providers. If the primary user (or delegator, respectively) wants to delegate access rights to any of his backups he is informed that he can revoke these access rights anytime in the future. He is also informed that in case that any delegate fulfils all conditions for accessing primary user's backup data, this data can be irrevocably copied and stored by the delegate. The primary user (delegator respectively) is asked if he understands possible risks and consequences of the delegation and if he really wants to delegate selected access rights which enable selected third parties to perform selected operations on the selected backup data in case that selected access conditions are satisfied. Delegation of access rights proceeds after the primary user (delegator, respectively) confirms it.

Third parties are becoming delegates of a particular backup data if they accept a delegation request. Before accepting, every delegate candidate is informed what partial identity initiated the delegation request, for what data, and under which conditions will the delegator be able to perform what operations on this data. The delegate candidate is informed that he can refuse (revoke) delegated access rights anytime in the future if he accepts it.

The storage provider should be able to cancel the contract in case that primary user (backuper, respectively) violates conditions of storage provider. In the demonstrator there is an interface which informs primary user (backuper, respectively) that storage provider cancelled the contract due to specific reason and gives primary user (backuper, respectively) advice how to reallocate backup data.

Integration in the demonstrator. These solutions will be addressed in the form of conceptual specification in the demonstrator.

Solutions for transparency on linkage and linkability

A primary user must be aware of potential risks regarding linkage and linkability of his actions, data, areas of life and others (see Section 3.1.1 for details) when operating with the demonstrator and he must be provided adequate information on how to avoid this risks.

Solutions for the demonstrator. In our demonstrator this requirement will be solved by informing the primary user (delegator, respectively) that his areas of life or partial identities will be linked together if selected access rights will be delegated to selected delegate candidate. The demonstrator will also inform primary user (delegator, respectively) under which conditions linkage will occur in case that delegate candidate receives access rights. Delegation request will be sent to delegate candidate only when primary

user (delegator, respectively) confirms that he is aware of potential risk of linkage of his areas of life (or partial identities).

Integration in the demonstrator. This solution will be directly implemented in the demonstrator.

Solutions for Privacy and Security Breach Notification

In case of a security breach of some security mechanism integrated to the demonstrator or in case of security incident of any storage provider the demonstrator derives benefit from, there should be some mechanism which informs the user about this incident and possibly gives him advice on how to cope with it (see Section 3.1.1 for details).

Solutions for the demonstrator. In our demonstrator first requirement will be solved by detection mechanism which monitors possible breaches of security functions utilized by the demonstrator. Second requirement will be solved by communication mechanism which informs primary user (backuper, respectively) about security incidents of the storage providers providing remote storage space to him. Additionally both of these mechanisms will provide information what are the possible consequences and how to deal with them if possible.

Integration in the demonstrator. These solutions will be addressed in the form of conceptual specification in the demonstrator.

4.4.2 Solutions for Data Minimisation Requirements

This section introduces solutions which fulfil requirements on data minimisation introduced in Section 3.1.2. The general goal of the "data minimisation" is to minimise the risk of misuse of the data.

Solutions for Data Minimisation by Anonymisation and Pseudonymisation

Minimisation of linkability and observability of the primary user's (backuper's, respectively) actions should be achieved using diverse identifiers for different storage providers. This requirement should assure that different storage providers are not able to link his actions or data in case that they would be controlled by a single entity.

Solutions for the demonstrator. In our demonstrator this requirement will be achieved by generating new credential for each different storage provider and for each different backup (possibly utilizing anonymous credentials). Our demonstrator will also provide functionality for automatic generation of new credentials by using a generator of cryptographically secure pseudo-random number.

Integration in the demonstrator. This solution will be directly implemented in the demonstrator.

Additionally, any delegate should implicitly not be able to link or observe actions of the primary user belonging to different areas of his life or covered by his different partial identities.

Solutions for the demonstrator. This requirement will be solved by utilizing anonymisation service. Actions of the primary user will therefore stay unlinkable and unobservable among different areas of life or partial identities of the primary user for any authorised delegate. Anonymisation service will assure that the real location of the user cannot be traced by the delegator as well as by storage provider or potential attacker.

Integration in the demonstrator. This solution will be directly implemented in the demonstrator.

Further, any potential attacker is not able to observe or link actions of the primary user (backuper, respectively).

Solutions for the demonstrator. Also this requirement will be solved by utilizing anonymisation service. This will assure, that any attacker observing the communication between primary user (backuper, respectively) will not be able to find out that any two datagrams originated from that particular primary user (backuper, respectively).

Integration in the demonstrator. This solution will be directly implemented in the demonstrator.

Solutions for minimisation of storage of sensitive data

The storage providers as well as delegates should have only minimal access to personal and sensitive data of the primary user. As far as for the purpose of the demonstrator there is no reason for the storage provider to access backup data of the primary user the data should be confidential for the storage provider. The same holds for any other third party which is not explicitly authorised by the primary user (delegator, respectively) to access the backup data.

Solutions for the demonstrator. As far as confidentiality is required, our demonstrator will utilize encryption mechanisms so that confidentiality of the primary user's data is assured. In addition no unauthorised third party, including storage provider, can access the backup data.

Integration in the demonstrator. This solution will be directly implemented in the demonstrator.

Minimisation of storage of sensitive data principle applies also to delegates. Primary user (delegator, respectively) should have possibility to delegate the smallest possible amount of data respecting his areas of life and partial identities sufficient for the specific purpose of the delegation. Primary user (delegator, respectively) should also be implicitly warned that his potentially sensitive data might be irrevocably revealed to selected delegate candidates.

Solutions for the demonstrator. Confidentiality of the backup data will be solved implicit by data encryption before sending it to the remote storage managed by the storage provider. Delegation of sufficient access rights will be solved by secure access control mechanism. Selection of proper access rights will be under full control of the primary user (delegator, respectively). Primary user's data will be separated according to his different areas of life or partial identities such that distributed storage capacity is effectively utilized. Demonstrator will assist primary user (delegator, respectively) in selection of the proper access condition. When delegating access rights, demonstrator will also provide information about possible risks for that particular type of delegation.

Integration in the demonstrator.

- Encryption of the data before transfer will be fully implemented in the demonstrator.
- Delegation of sufficient access rights by secure access control mechanism will be implemented to such an extent which sufficiently demonstrates this functionality.
- Separation of primary user's data according to his different areas of life or partial identities will be implemented to such an extent which sufficiently demonstrates this functionality.
- Assistance of the demonstrator will be addressed in the form of conceptual specification in the demonstrator possibly with conceptual demonstrative implementation.
- Warning window will be directly implemented in the demonstrator.

Solutions for active support for data minimisation

Data minimisation should be actively supported by the demonstrator.

Solutions for the demonstrator. In our demonstrator this requirement will be solved by supporting security mechanisms, which assure unobservability and unlinkability of the primary user's actions as well as anonymisation and pseudonymisation of the primary user's identity.

Integration in the demonstrator. This solution will be directly implemented in the demonstrator to such an extent, which sufficiently demonstrates this functionality.

Solutions for Minimisation of the Time Frame of Data Exposition

The time frame of the access rights delegated to legitimate delegates should be limited to such a minimal extent, which is sufficient for the purpose of the delegation.

Solutions for the demonstrator. When delegating access rights, primary user (delegator, respectively) will be implicitly offered delegation valid within limited time frame only according to the purpose of the delegation. Demonstrator will provide the possibility to customize the range of the time frame. The primary user (delegator, respectively) will be warned in case that the time frame selected by user is too extensive according to the purpose of the delegation. He will be asked for explicit confirmation in case of delegating access rights which do not expire at all (with expiration set to infinity).

Integration in the demonstrator. This solution will be directly implemented in the demonstrator to such an extent, which sufficiently demonstrates this functionality.

Solutions for Minimisation of the Disclosure of Personal Data

For the practical solutions of the demonstrator, it is required that the primary user (i.e., delegator or backuper) should minimise the disclosure of his personal data.

Solutions for the demonstrator. When delegating access rights to delegate candidates, the primary user (delegator, respectively) is warned that as soon as the delegates gain access to his data, the primary user (delegator, respectively) has to rely on the trustworthiness of the delegates because in fact he has no longer direct control on what happens to his data then.

Risk of disclosure of personal data contained in backups will be solved by encryption of the data. Disclosure of personal data resulting from the relationship between storage providers and primary user (backuper, respectively) will be solved by utilizing anonymous (pseudonymous) credentials, anonymous payment system and by generating new identifier for each backup. Also communication between primary user (backuper, respectively) and storage providers will be anonymised. Communication between primary user (delegator, respectively) and delegates will be pseudonymous by default and without the possibility to detect their locations each other.

Integration in the demonstrator. Utilization of data encryption will be directly implemented in the demonstrator. The further above-mentioned solutions will be addressed to such an extent, which sufficiently demonstrates their functionality.

Solutions for Minimisation of the Linkability and Linkage of Personal Data

For the implementation of the demonstrator, the primary user (i.e. delegator or backuper) should minimise the linkability and linkage of his actions and data, especially among different areas of his life or partial identities.

Solutions for the demonstrator. Linkability of the data and actions of primary user (delegator or backuper, respectively) according to delegates will be avoided by utilizing anonymisation service separating different areas of primary user's life and his partial identities. Linkability of the data potentially collected by storage providers, attackers or other third parties (legally related or not) will be address by integrating anonymisation functionality, anonymous (pseudonymous) credentials and by using generating unique credential for each backup.

Integration in the demonstrator. Demonstrator will address the above-mentioned solutions to such an extent, which will demonstrate this functionality in sufficient manner.

Solutions for Minimisation of Multipurpose or Context-Spanning Use of Data

The multi-purpose or context-spanning use of data should be minimised.

Solutions for the demonstrator. This requirement will be solved in the demonstrator by utilizing anonymisation service. Moreover, anonymous (pseudonymous) credentials unique for every backup will be used in the demonstrator in order to assure context separation.

Integration in the demonstrator. Demonstrator will address these solutions to such an extent, which sufficiently demonstrates this functionality.

Solutions for Data Minimisation by Unique Identifiers

Data minimisation should be achieved by using unique identifiers which may be used in different contexts.

Solutions for the demonstrator. This requirement will be solved by generating new identifiers for every backup stored in distributed environment in the storage space provided by storage providers.

Integration in the demonstrator. The above-mentioned solution will be directly implemented in the demonstrator to such an extent, which demonstrates this functionality in sufficient manner.

Solutions for Data Minimisation by Anonymous or Pseudonymous Authorisation and Access Control

The actions between the primary user (backuper or delegator) and the storage providers or between the primary user and the delegates as well as delegates and storage providers should be supported by anonymous or pseudonymous authorisation and access control.

Solutions for the demonstrator. This requirement will be solved by utilizing anonymous or pseudonymous credentials mechanism between all above-mentioned parties.

Integration in the demonstrator. This solution will be partially implemented in the demonstrator to such an extent, which demonstrates this functionality in sufficient manner. Part of this solution will be addressed in written form as conceptual specification in the demonstrator.

Solutions for Data Minimisation by Minimising Irrevocable Consequences

The primary user (delegator, respectively) should always be able to revoke access rights delegated to delegates.

Solutions for the demonstrator. This requirement will be solved by integrating a mechanism which allows the primary user (delegator, respectively) to revoke his access rights delegated to the corresponding delegates. This action must also generate message to corresponding delegates that their rights have been removed by the delegator.

Integration in the demonstrator. This solution will be implemented in the demonstrator in such a way, that it sufficiently demonstrates above-mentioned functionality.

Irrevocable consequences concerning the privacy of the primary user (backuper, respectively) should be minimised irrespective of his ability to manage his privacy, i.e., it should be possible to remove any backup item contained in any backup he previously created.

Solutions for the demonstrator. This requirement will be solved by implementing a function that allows a primary user (backuper, respectively) to delete any backup item and all instances of it in whichever backup created by the primary user.

Integration in the demonstrator. An according solution of this requirement will be fully implemented in the demonstrator.

4.4.3 Solutions for Privacy-Related Requirements Derived from the Backup and Synchronization Nature of the Demonstrator

This section introduces solutions which fulfil privacy-related requirements derived from the specific nature of the demonstrator introduced in Section 4.3.3.

Solutions for Localization of the Backup Data

Section 4.3.3 requires that the primary user and the authorised delegates should have a mechanism which allows them to visualize what kind of data is stored in which backup with the possibility to search the data according to selected properties. In addition, there should be a search functionality which allows them to utilize search requests based on several search attributes. On the top of that, all of these actions must be performed in a secure manner so that no attacker is able to reveal what backup item is/was searched, who was searching it or even that it was searched.

Solutions for the demonstrator. This requirements will be solved by a mechanism which visualizes in which backups (and in how many copies) are particular backup items stored and in what state (e.g., time of last update, older archival version, time of last synchronization). The secure search functionality will utilize anonymisation mechanism in order to avoid linkability and observability. The search functionality will primarily access information about the backup items structure stored locally in order to avoid communication overhead. Information about the structure of the backup items will be stored in special area of each backup. Corresponding secure synchronization of the information about the structure of the backup items will be assured by a special mechanism in anonymous, unlinkable and unobservable way.

Integration in the demonstrator. These solutions will be implemented in the demonstrator to such an extent, which sufficiently demonstrates the above-mentioned aspects.

Solutions for Backup and Removal of a Single Item

The primary user should be able to insert respectively delete any single item to respectively from any of his backups (see Section 4.3.3 for details).

Solutions for the demonstrator. Apart from the need to encrypt the data before sending it to the backup (see Section 4.4.2), this requirement results in the need to utilize an encryption schema which allows to insert or remove encrypted data items. This needs to be performed in a secure manner with respect to anonymity, unlinkability and unobservability.

Integration in the demonstrator. A mechanism which fulfils all of the above mentioned properties will be implemented in the demonstrator to such an extent, which sufficiently demonstrates the required functionality.

Solutions for Back-In-Time Recovery

The primary user should be able to use a back-in-time functionality which will allow him to recover not only the most recent version of the backup item created during the last back up action, but also previous versions containing previous state of the corresponding primary item which was backed up in the past (see Section 4.3.3 for details).

Solutions for the demonstrator. A mechanism which provides back-in-time backup and recovery functionality will be provided by the demonstrator.

Integration in the demonstrator. This solution will be addressed in the form of conceptual specification in the demonstrator.

Solutions for Full Deletion

Section 4.3.3 requires that the primary user (deleter, respectively) should be able to perform full deletion of any of his backup items (even all backup items), including its

copies and older versions, distributed in different backups stored on storages of different storage providers in secure manner supporting revocability of the storage of the data.

Solutions for the demonstrator. An appropriate mechanism which performs full deletion will be integrated in the demonstrator.

Integration in the demonstrator. This solution will be addressed in the form of conceptual specification in the demonstrator.

Solutions for Backup Recovery after Unrecoverable Crash of the User's System

The primary user should be able to recover all of his backup data (stored in backups) even if he would permanently lost access to his system and data on it (including the demonstrator installed on it) (see Section 4.3.3 for details).

Solutions for the demonstrator. This requirement will be solved by implementing a mechanism which will allow the primary user to create secure backup of his credentials used for accessing services of storage providers. This mechanism will allow user to export his credentials to secure media and import them to the demonstrator again in case of system crash.

Integration in the demonstrator. This solution will be implemented in the demonstrator to such an extent, which sufficiently demonstrates the required functionality.

4.5 Further Potential Scenarios and Use Cases

This Section deals with further potential scenarios and use cases that may extend the current scope of the privacy-enhanced backup and synchronisation demonstrator.

The developers of the WP 1.3 demonstrator should consider how the scenarios in this chapter may influence the design of the privacy-enhanced backup and synchronisation demonstrator. Most technical relevance has probably the possibility to migrate the used cryptographic functionality to other schemes (see Section 4.5.2).

4.5.1 Handling of Incidents

There could be a variety of incidents that would have to be tackled in case of occurrence. This section picks two relevant areas, namely the violation of the contracts between entities involved as well as the issue of search and seizure.

Violation of the Contracts Between Entities Involved

The relationship between the primary user and the storage provider is determined by a contract (the End User License Agreement). This contract has to contain information about the rights and duties of the respective parties. In particular the storage provider usually gives information on the planned or guaranteed availability of its service. In

addition, a potential payment by the user for the service may be laid down in the contract. In case of violation of the contract by one party, the other party usually can cancel the contract. Here it is important to think about the consequences regarding the data that is being stored and the expectation of the primary user or potential delegates that the backup items are accessible.

When drafting or entering such a contract, it should be clear from the beginning:

- How can it be ensured that the primary user will maintain the control of personal data stored at the storage provider if the contractual relationship between the primary user and the storage provider is terminated? How long will the data be kept? And as soon as the data are deleted: Will the data be safely erased? If the data are still there: Will access of the primary user or a delegate be denied?
- In case of a paid backup and synchronisation service: What processes are established if the primary user is late with the payment or stops to pay for the service? For instance, if a regular payment is part of the contract, but the primary user is in hospital in need of the delegation functionality, but at least temporarily without the possibility to initiate the payment, how could this be handled? Clearly the primary user then should be informed about the missing payment, but should also assigned delegates be informed? Under which conditions should the storage provider reveal the name and address of a pseudonymous primary user to collect her debt? And what does this mean for the accessibility of the data – should the storage provider have the possibility to prevent any access of primary users and delegates in case of lacking payment? This may have critical effects.
- In case of bankrupt, mergers or sales of corporations on the storage provider's side: How can it be guaranteed that the level of protection originally ensured to the primary user will remain at least equivalent?

Also, a change of policies (privacy policies as well as general terms and conditions) on the storage provider's side may be a violation of the contract, but there may be changes that are in line with the contract, or there may be legally demanded changes that leave the storage provider with no option but to adapt the contract. In all these cases it would be necessary to inform the primary user and also delegates if the changes may affect them about the changes. Usually the persons involved have to affirm their consent.

In addition to the contract between the primary user and the storage provider there may be contracts between the primary user and the delegator (if they are not the same) or between the delegator and the delegate. Further the potential involvement of a delegation supervisor could be based on a contractual relationship. In all of these cases there may be questions of payment or liability issues if an entity does not act according to the predefined rules. How to handle these possible incidents should be clarified in advance to prevent any unpleasant surprise in the future. In general, breaches of confidentiality, integrity or availability guarantees concerning the backup items or credentials should not happen, but if they occur, the other parties involved should be informed about the incident and possible precautions they can take to minimise undesired consequences. This can happen on the storage provider's side, on the primary user's side or on a delegate's side. There also may be liability issues, and compensation rules may be foreseen in the contract. Otherwise this may justify a legal claim for damages.

Effects of Search and Seizure

There is a special case, which may have to be treated differently: search and seizure. In case police or law enforcement suspect that a crime has been committed, they have according to many civil law and common law legal systems the right to do a search of a person's property and confiscate any evidence that is potentially relevant to the crime. Regarding the backup and synchronisation system this may in principle happen on the side of all entities involved, i.e., the storage provider, the primary user, the delegator, the delegate, the credential issuer.

Search and seizure can mean that only a copy of the data from the IT systems are taken, but they are still functioning, or it can mean that the IT systems are taken away, often for some days or several months. In the latter case there is probably at least a downtime of some time until the functionality of the IT systems involved can be restored. The availability of the data and the functionality of the systems can be reduced even more if restoring the data is not possible.

Possible consequences can be that the service does not work or that parties such as the delegate cannot fulfil their tasks any more in case their credentials are gone. The primary user whose backup items are at stake is not necessarily the suspect in this scenario, but still may be influenced by a search and seizure procedure.

Depending on the legal scheme and the exact wording of a court order or other documents that have to be shown in a search and seizure procedure, it may be allowed or not allowed to immediately inform other parties about this incident. All parties involved should at least document what is happening when, so that later on (e.g., when a suspect has been cleared) the incident and possible consequences can be reconstructed. Usually the authority in charge has to inform the suspected individuals at least afterwards about the search and seizure procedure, even if they didn't notice it and no charges are pressed.

4.5.2 Handling of Technical Changes

Progress in technology provides new opportunities, but also poses new challenges. This section exemplifies that by the possibility of migrating the cryptographic functions as well as the deployment of cloud computing technology.

Possibility of Migrating the Cryptographic Functions

Long-term protection of privacy and security poses various challenges. One of these challenges is how to maintain a high level of protection by cryptographic means. It is foreseeable that today's assumed strength of cryptographic modules will not be kept for a period of several years. Instead, it will be necessary to migrate to new cryptographic algorithms or other safeguards.

This plays an important role in the privacy-enhanced backup and synchronisation system since cryptographic modules will be key components. Here the concept of the demonstrator should elaborate how a migration of the cryptographic functionality is possible and how parties involved will be informed about the necessity of migration including the potential consequences when migrating or not migrating that functionality.

Further it could be discussed whether for the sake of robustness different cryptographic means should be used in parallel. Hence, in case one of the algorithms or

implementations has to be considered not safe any more, the other one would still ensure a high level of protection. However, implementing this may be too sophisticated for the purpose of a demonstrator.

Storing Backup Items in a Cloud

In the current concept of the WP 1.3 demonstrator, the set of storage providers seems to be clearly defined in advance. According to the concept, the primary user can always be aware of where the backup items are located.

A different setting would be the storage of backup items in a cloud. A definition of cloud computing is provided by the National Institute of Standards and Technology (NIST) : "Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."

Cloud computing poses many legal challenges, especially on the issue of responsibility for the (personal) data, liability issues, the relationship between different members of the cloud (who is a data controller, who is a data processor, which contracts exist?), the applicable law and questions concerning possibly trans-border data flow. Surely it is not an unrealistic scenario, but for the demonstrator it should not be tackled because this would unnecessarily complicate its design.

4.5.3 Scenarios to Support Users

The backup and synchronisation system probably will require primary users to explicitly configure which data belongs to which partial identities and the exact delegates. There might be more convenient solutions that would require supporting mechanisms. This section describes two of those possible mechanisms.

Support Primary Users in Sorting their Data

Handling several partial identities for various areas of life will not be that easy for users. This will be even more difficult or at least cumbersome if they have to think of different delegates for separate areas of life. A manual configuration via the backup and synchronisation system would not be very convenient.

However, it would be possible for communication partners of the primary users including the issuers of official documents such as diploma certificates, school reports or tax IDs to attach information on how to keep them. This attached information may comprise how confidential they should be stored, how quickly they should be available, who should be able to access them under which conditions, with whom to share, how to involve delegates and inform them about their tasks concerning specific data items etc. A standardised set of meta-information that is known and interpreted by the privacy-enhanced backup and synchronisation system could be very helpful for users to sort their data and handle them appropriately.

Third-Parties Identified by Role

The primary user could choose to grant access to individuals representing a role, without necessarily be identified by individual information. This could be defined as a group of several persons playing the same role, defined, e.g., by a particular position in a given organisation. It might also be automatic that the primary user, by taking a position in this organisation, delegates without having to express this delegation specifically, i.e. her own role conveys the delegation to any person fulfilling another role concerning backup items of specific partial identities. Still the user interface should clearly communicate who is able to get access.

In that matter, ontologies could be used to define classes or equivalences of roles. Then the organisation policies could include rules that would use those classes to determine the access control rights for the primary users. Again this would support users in handling their data and delegates in the backup and synchronisation system.

4.6 Conclusion

This section presented the concept of the privacy-enhanced backup and synchronization demonstrator selected as the basis for the third year's focal demonstrator. It clarified what the requirements on the demonstrator are. It also outlines the proceeding how they will be fulfilled in the WP 1.3 focal demonstrator. It was presented that the objectives of lifelong privacy lead to practical results, which can be applied for solving real-life issues in enhanced ways. Our demonstrator reveals new problems, which emerge as soon as lifelong aspects related to the data subject are taken into consideration. We presented a new approach, which can help an average citizen to protect himself against unwanted data loss respecting his different partial identities and areas of life. Our approach proceeds in such a way that it takes into account lifelong aspects of a human being and corresponding implications within the scope of privacy. Furthermore in this document, we also clarify the reasons which led us to the decision for selecting the backup and synchronization area as the fundamental base of our further focus.

With this concept the demonstrator directly addresses several issues of "Privacy for Life": It provides a possibility for users to store their data safely over a long period of time, it distinguishes between various areas of life by separating the data of different partial identities, and it offers delegation of access rights which may be necessary if the user cannot manage her (backup) data on her own.

Chapter 5

Conclusion

5.1 Lessons Learned

The aim of this deliverable was to analyse the research field of Privacy-Enhancing Identity Management in the setting of enabling *lifelong* privacy.

In this regard, we identified not only particular aspects to be considered when examining and developing for the given area – such as stages of life, areas of life, identities and partial identities etc. – but we were also reflecting on requirements imposed by the specifics of lifelong privacy management. Those requirements have been looked at from different viewpoints: on a high-level basis, from social-cultural interests and related to delegation, which is one of the interesting peculiarities of lifelong privacy with respect to the ability to manage one's privacy throughout his life.

In particular, the requirements elaborated regard issues of

- transparency including openness, notice, awareness, and understanding;
- data minimisation;
- controllable and controlled data processing;
- user-controlled identity management;
- delegation;
- practicability of mechanisms;
- change management.

Besides the definition of those stipulations rather general for the given research field, further requirements have been focusing particularly the actual nature of the chosen prototype, which is a backup and synchronisation demonstrator respecting lifetime aspects. The selection of this demonstrator is accordingly motivated within this deliverable and opens up interesting new challenges with respect to lifelong privacy and privacy management.

Accordingly, this document lists corresponding solutions to the indicated requirements applied to the actual demonstrator.

Obviously, not all requirements indicated in this deliverable are possible or even not wanted to be realised solely on a technological basis. There are further areas that are more suited to solve issues of the requirements in some cases. Also, economic contemplations may play an important role. Thereby, approaching the problem field on an economic basis may imply monetary considerations, on the one hand, but also finding an appropriate trade-off between getting service and retaining privacy, on the other hand.

Since economic solutions are not considered within the PrimeLife project, this might be an interesting topic of a follow-up project.

Apart from the economic solutions, some of the requirements demand legal regulations. A selection of the most important requirements to be handled by policy makers have been identified and are described in the following section.

5.2 Recommendations for Policy Makers

Many of the requirements mentioned in Chapters 3 and 4 do concern issues that are also relevant for policy makers. As stated in those previous chapters, existing regulations are often not sufficient to guarantee proper handling of personal data throughout life for the data subject. The law may be interpreted in different ways and in many situations there are no legal stipulations for the orientation in handling personal data.

Note that in our understanding the term "policy makers" comprises not only law makers setting legal standards, but also those entities that set standards by other means, in particular by standardisation of technology, but also by presenting best practices as guidelines for system developers or application providers. However, the issues pointed out in the following sections cannot be sufficiently handled only by technological solutions, so our recommendations should be reflected at least in the legal systems within Europe.

5.2.1 Openness, Transparency, Notice, Awareness, Understanding

To ensure transparency, policy makers need to clearly consider all consequences when deciding about new statutes regarding personal data. Especially revocation needs to be considered in each decision making process. Therefore policy makers have to take into account which situations may occur where revocation of personal data is necessary and which risks and effects new statues on personal data may have. One example may be identification numbers that are assigned to natural persons and where legislators have to clearly define on the revocability and to consider effects and risks of the ID number (for example, the taxID introduced in Germany in 2008 as unique identifier for each citizen where personal data may be accessible for unauthorised parties).

Transp-Req m): Policy makers should define and explicate areas where revocable respectively irrevocable consequences are demanded respectively prohibited. This should guide system developers when conceptualising, designing and implementing ICT systems as well as application providers when operating applications.

In many situations throughout life the interpretation of law is various. Often definitions are unclear and leave room for all ways of interpretation or are not complete. Therefore, policy makers and supervisory authorities should make clear, what they demand from data controllers and data processors concerning privacy-relevant data processing, for example, how to interpret privacy regulations.

This problem, for example rises in the context of joint responsibilities for data processing. In many situations personal data may be relating to individuals but also to groups of individuals. Therefore the discussion raises how to handle personal data in case of joint responsibility of personal data. In general there is no "ownership" of data. Personal data relate to an identified or identifiable natural person, but is not a property. For this reason the wording ownership" is not correct and a better way to name the situation would be joint responsibility. Every data subject is responsible for the data within his or her area of accountability. Therefore, for example it has to be asked for the consent if personal data of a third person should be processed (for example, posting of pictures in SNS).

Transp-Req n): Policy makers and supervisory authorities should make clear what they demand from data controllers and data processors concerning privacy-relevant data processing, i.e., how to interpret privacy regulation.

5.2.2 Decreasing the risk to Privacy Throughout Life by Data Minimisation

The concept of "pseudonymous convertible credentials" (cf. Section 3.2.3) provides privacy-enhancing ways to combine anonymity and accountability requirements in ICT systems: They enable a data subject to prove her authorisation whilst controlling the conditions determining her identifiability and accountability at the same time. However, to foster their employment in ICT systems, an appropriate infrastructure has to be built up, and the concept has to be reflected in legal provisions [RBB+08].

DatMin-Req l): Policy makers should support setting up and standardising the necessary infrastructure for issuing pseudonymous convertible credentials and their usage in their ICT systems where appropriate.

DatMin-Req m): Policy makers should evaluate current legal provisions in the light of pseudonymous convertible credentials.

5.2.3 Controllable and controlled data processing

This section lists a few recommendations concerning the legal provisions on data processing, the question of user control, sanctioning privacy infringements, conflicting policies, and delegation.

Real purpose binding. Considering long-term effects, some of the currently existing legal provisions could need some intensification. This applies to the principle of purpose binding which has been weakened by manifold exceptions as well as to handling of sensitive data:

> *Control-Req o):* Policy makers should further limit exemptions to use (potentially) personal data for other purposes.

> *Control-Req p):* Policy makers should be extra cautious with (potentially) sensitive data.

User control. The data subject is assumed to be an autonomous, privacy-aware individual, capable to act appropriately according to her own will. From the previous sections, it is clear that this ideal image is not true for all cases. This also affects principles such as "consent" which is a sufficient basis for data controllers to process personal data. But if the data subject cannot understand what she is consenting to, the consent is obviously not a useful solution. And in today's complex world this may become the rule.

> *Control-Req q):* Policy makers should rethink the concept of consent and possibly limit data processing based on consent in its scope or extent.

Hence, the whole concept of user control has to be challenged: It should not be mistreated to shift the responsibility on to an overburdened individual, but users should be empowered so that they really can exercise their rights.

> *Control-Req r):* Policy makers should seek for ways to efficiently implement fair user control, easy to perform for all individuals.

Coping with privacy infringements. Most of the data protection laws already have regulations regarding sanctions for privacy infringements. Nevertheless they are often not applicable or enforceable and furthermore differentiate within the EU Member States. Therefore useful legal sanctions for responsible parties as well as remedies for victims should be legally stipulated.

> *Control-Req s):* Policy makers should revise the current framework for sanctioning privacy infringements and providing remedies for victims.

> *Control-Req t):* Supervisory authorities should sanction privacy infringements by noticeable punishments.

> *Control-Req u):* Policy makers should elaborate concepts for achieving remedy for victims of privacy infringements ("privacy infringement insurance"?).

Dealing with conflicting policies and multiple processors. When, for example, the data subject and the controller use different policies, the problem may appear that there are conflicts within the policies. There might be different preferences for data processing within the policies of the user and the controller. From the legal aspect, the data subject may define the policy for the handling of her personal data. If, for example, the controller wants to use personal data for different purposes than defined in the data subject's policy, he does need the explicit consent of the data subject. That means if a policy should be changed with regard to the purpose of the data processing, the consent is required.

From the technical point of view, it might be helpful to define structures for policies. Policies could have clearly defined contents that can not be changed by the parties involved. Other parts of the policy may be changed by the parties involved in a clearly defined way. One solution can be joint policy responsibility.

In general, policies can only be evolved for the future up to a certain level. But it also has to be possible to adjust policies to current technical and legal changes. Therefore mechanisms are needed that enhance and adjust policies to current changes.

This also raises the question what happens with the revised policy and if transitional periods have to be considered. To keep policies up to date there could be, for example, a yearly reminder for the data subject or the controller for the decision about a need of policy change. At the moment there are no processes and no transparency for the data subject.

Furthermore, there might appear a conflict if multiple processors are involved in data handling. Processors handle personal data on behalf of the controller and merely have an auxiliary function with regard to the processing.[1] With regard to Art. 17 of the Data Protection Directive, the controller must choose a processor providing sufficient guarantees in respect of the technical security measures and organisational measures governing the processing to be carried out. The carrying out of processing by a processor must be governed by a contract that stipulates that the processor shall act only on instructions from the controller.[2] Processors do not have to comply with the vast majority of requirements determined by the Data Protection Directive, but basically must follow the instructions of the controller and implement appropriate technical and organisational measures ensuring data security.

> *Control-Req v):* Policy makers should propose guidelines for how to deal with conflicting policies.

Delegation. Delegation may not only be useful for cases where the concerned individual is fully in possession of his/her mental capabilities and decides on her own to transfer the exertion of rights to another person. Proxies often are overtrained with their duties or even do not know the limitations of their responsibilities. Therefore law

[1] Art. 2 (g) of the Directive 95/46/EC

[2] The European Committee for Standardisation (CEN) has published an Article 17 Model Contract (Standard form contract to assist compliance with obligations imposed by Art. 17 of the Directive). This contract is online available at: ftp://ftp.cenorm.be/PUBLIC/CWAs/e-Europe/DPP/CWA15292-00-2005-May.pdf.

makers should define general principles or guidelines for handling of privacy by a proxy as an orientation.

> *Delegate-Req l):* For legally relevant settings, policy makers should regulate the circumstances of expressing and revoking delegation.

> *Delegate-Req m):* Policy makers should define general principles or guidelines for the handling of privacy by a proxy as an orientation.

> *Delegate-Req n):* Policy makers should provide rerequisites to enable later revision of privacy-relevant actions performed by the proxy on behalf of a data subject.

> *Delegate-Req o):* Policy makers should provide reasonable legal solutions to protect the proxy and to balance the interests of the proxy and the principal in a fair way.

5.2.4 Change Management

In Section 3, the necessity of dealing with changes has already been elaborated – primarily from the perspective of a data controller, but also other stakeholders have been mentioned:

> *ChangeMng-Req b):* Data controllers, data processors, system developers, and policy makers should monitor changes in society, law, and technologies and react appropriately (for example, by evaluating chances and risks, adapting current processes, regulation or standards to the changed conditions).

Ensuring legal compliance over time. Not only data controllers, but also supervisory authorities dealing with privacy and data protection have to tackle changes in law, in the state-of-the-art of technology. This has to be reflected in their regular work, e.g, when controlling legal compliance of data processing by disposing audits or seals of approval.

Reacting to societal changes – legal and technical aspects. It also has to be taken into account, that with a changing society also legal and technical aspects may change. Here the question raises how jurisdiction and technology react on these changes and how in general these changes can be recognised. This could be on the one hand by the feedback of the society and it can be used to try to develop law and technologies compliant to the changes. Therefore changes have to be achieved by developers. Developers need feedback mechanisms to react on society changes with legal and technical implementations. To enable evaluation and feedback building safeguarding technologies may (deliberately) reduce degrees of freedom in action and freeze the current state of

policy. However, technology should not work against evolvement of societal consensus. Thus, solutions should take into account the possibility to evaluate whether the current state is still considered alright and provide for an optional feedback if changes are desired. For this feedback it is also important to consider privacy implications, for example, providing possibilities for individuals giving feedback to stay anonymous [Phi04].

> *ChangeMng-Req c):* Policy makers should establish (privacy-enhancing) feedback mechanisms for society concerning privacy-relevant changes (for example, on attitudes what is to be considered private or public).

> *ChangeMng-Req d):* Policy makers should establish (privacy-enhancing) feedback mechanisms for society concerning privacy-relevant regulation, processes and technical implementations.

Ex ante privacy assessment of technical advancement and legislation of emerging technologies. Today, development of privacy law is mostly one step behind technological developments. Amendment of law only responds to insights into the consequences of advances in the processing and analysis of personal data [Kir08]. Preventive approaches on revising privacy law in the light of expected advances in technology are not common, and developing high-level concepts or formal models for a systematic assessment of emerging technologies is still work in progress [CHP+09]. Dommering even emphasises: "When a technology reaches a vast diffusion it affects society in a way which was not part of the design" [Dom06].

Prior checking and technology assessment are two instruments which should enable more than a glimpse on what implications upcoming technologies or ICT systems may have concerning privacy. This should also be intensified in the law-making process. In particular the assessment should be done in an interdisciplinary approach comprising lawyers, technologist, sociologists, psychologists, economists, perhaps also philosophers, historians, or physicians. This enables a broader discussion on potential cross-effects and should yield a wider perspective on the possible implications.

> *ChangeMng-Req e):* Policy makers should demand and support ex ante privacy assessments of technical, regulatory, and legislatory advancements.

This also applies to standardisation activities or funding of projects.

> *ChangeMng-Req f):* Policy makers should consider the state-of-the-art and research results concerning privacy-enhancing technologies as well as potentially privacy-infringing technologies in law making, standardisation, funding and other supporting actions.

This chapter shows that – especially for policy makers – it is quite a challenging task to enable identity management throughout life. This discussion of recommendations is

not exhaustive and only addresses main problems in selected scenarios. But in general, policy makers need to adjust the existing privacy policies to current social and technical development.

Glossary

Area of life (AoL)

sufficiently distinct domain of social interaction that fulfils a particular purpose (for the data subject) or function (for society).

Attacker

is an entity, which performs malicious activities trying to violate mechanisms of the privacy-enhanced backup and synchronization demonstrator. These are mostly activities, which can lead to the achievement of unauthorised access rights to the backup data, which can cause damage or unauthorised modification of the backup data or unauthorised damage or modification of the relation between the backup and the involved third party (mostly primary user). Further activities of the attacker also cover unauthorised linkage of the backup data, which was intentionally separated (physically or/and logically), or unauthorised linkage of the actions of the primary user or involved third parties performed on the backup. Last but not least, activities of the attacker are also those actions which have unwanted impact on the mechanism of conditional access control and which are not authorised by the primary user.

Backup

a non-empty set of backup items.

Backup item

a copy of a primary item stored in the backup. A backup item reflects the data of a primary item in the time when the backup item is created. Note that even if each backup item must belong to one and only one primary item, this primary item may not exist during the whole lifetime of the backup item. A backup item can exist in several versions in a particular point of time. The previous versions of a backup item backed up in the past are called predecessors of a backup item. Any version which is older than the current version of a backup item is considered to be its predecessor. Future versions of a backup item which will be created in the future are called successors of a backup item. All versions of a backup item which are created after the current version are considered to be its successors. Current version of a backup item is the last version which exists in the current time.

Backup recovery/restoration

the process of extracting an original primary item from a corresponding backup item, which was previously created during the back up process. The outgoing

primary item gained from a backup recovery/restore process has the same state as the previous state of that primary item when the back up process was performed on it.

Backuper

initiates the back up action. In most applications of this demonstrator, the primary user acts as the backuper.

Credential issuer

is an entity, which issues a credential verifying a certain status of the primary user. This status can for example be: "primary user is ill", "primary user is hospitalized", "primary user is dead" or others. A credential issuer must be authorised by a corresponding authority (e.g., governmental) for issuing a certain type of credentials.

Data controller:

the natural or legal person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing of personal data; where the purposes and means of processing are determined by National or Community laws or regulations, the controller or the specific criteria for his nomination may be designated by national or Community law [Eur95, Art. 2d].

Data processing:

any operation or set of operations which is performed upon personal data, whether or not by automatic means, such as collection, recording, organisation, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction.

Data processor:

a natural or legal person, public authority, agency or any other body which processes personal data on behalf of the controller [Eur95, Art. 2e]

Data subject:

an identifiable natural person, which is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity [Eur95, Art. 2a].

Delegate

is an entity, which receives particular rights on the backup from a delegator.

Delegate candidate

is an entity, which was selected by delegator to act like delegate but does not possess particular rights yet.

Delegation request

is a request sent to the delegate candidate asking him whether he accepts particular rights from the delegator.

Delegation:

a process whereby a *proxy* (also called delegatee or agent) is authorised to act on behalf of a *principal* (also called delegator) via a mandate, i.e., transferred duties, rights and the required authority, from the principal to the proxy.

Delegator

is an entity, which has the privilege to delegate rights to delegates concerning a particular backup. In most applications of this demonstrator, the primary user acts as the delegator.

Deleter

initiates the delete action on a particular backup. In most applications of this demonstrator, the primary user acts as the deleter.

Full lifespan

the range of time from the emergence of the first information that is related to the human being (from the moment of birth until the death of the data subject) until the point in time when no more personal data is generated.

Legally related party

is anyone being in a legal relationship with the primary user or the storage provider.

Personal data:

any information related to an identified or identifiable natural person. Natural persons are only living individuals but neither deceased nor legal persons [Eur95, Art. 2a].

Primary item

an original item for which one or more backup items are created during the back up action. In a general sense, a primary item can be referred to as any determinate set of data, which has one or more copies called backup items dedicated for backup purposes. A primary item can be a file but it can also be a more specific type of data as for instance an e-mail, a contact, or even settings of the TV.

Primary user

data subject who owns/holds primary items.

Restorer

participates on the backup recovery/restoration action and obtains the content stored in a particular backup as the result of successful backup recovery/restoration action.

Service provider:

a natural or legal person that operates an application based on an ICT system and offers it to users.

Stage of life (SoL)

a stage of life of an individual with respect to handling his privacy is a period in the life of this individual in which the ability to manage his private sphere remains between defined boundaries characterizing the current stage of his life [CHP+09].

State of life

temporary or permanent state of the data subject's life, which can be certified by a corresponding credential issuer and which might have impact on the ability of the data subject to manage his data (e.g., illness, hospitalization, death, pregnancy, imprisonment and others).

Storage

physical or logical device providing storage space for the backups of the primary user.

Storage area

destination where the storage is located. In our demonstrator this is mostly remote location administered by a particular storage provider accessible to the primary user and delegates virtually via communication network.

Storage provider

provides storage space for backups.

To back up

the process of creating copies of the primary item into one or more backup items and storing them in a corresponding storage.

User

User means any natural person using a publicly available electronic communications service, without necessarily having subscribed to this service [Eur02, Art. 2a]

Bibliography

[AB07] Rafael Accorsi and Matthias Bernauer. On privacy evidence for UbiComp environments – Broadening the notion of control to improve user acceptance. In *Proceedings of the 5th Workshop on Privacy in UbiComp*, pages 433–438, 2007.

[ABB+09] Claudio Ardagna, Carine Bournez, Laurent Bussard, Michele Bezzi, Jan Camenisch, Aleksandra Kuczerawy, Sebastian Meissner, Gregory Neven, Stefano Paraboschi, Eros Pedrini, Ulrich Pinsdorf, Franz-Stefan Preiss, Slim Trabelsi, Christina Tziviskou, Pierangela Samarati, Jan Schallaboeck, Stuart Short, Dieter Sommer, Thomas Roessler, Sabrina de Capitani di Vimercati, Mario Verdicchio, and Rigo Wenning. Final Requirements and State-of-the-Art for Next Generation Policies. Primelife Project Deliverable D5.1.1, PrimeLife Project, Aug. 2009.

[AG05] Alessandro Acquisti and Jens Grossklags. Privacy and Rationality in Individual Decision Making. *IEEE Security & Privacy*, 3(1):26–33, Jan. 2005.

[Art03] Article 29 Data Protection Working Party. Working Document on biometrics of 1 August 2003. http://ec.europa.eu/justice_home/fsj/privacy/docs/wpdocs/2003/wp80_en.pdf, 2003.

[Art05] Article 29 Data Protection Working Party. Working Document on a common interpretation of Article 26(1) of the Directive 95/46/EC of 24 October 1995. http://ec.europa.eu/justice_home/fsj/privacy/docs/wpdocs/2005/wp114_en.pdf, 2005.

[Art07] Article 29 Data Protection Working Party. Opinion 4/2007 on the concept of personal data. 01248/07/EN, Working Paper 136. http://ec.europa.eu/justice_home/fsj/privacy/docs/wpdocs/2007/wp136_en.pdf, 2007.

[Art08] Article 29 Data Protection Working Party. Working Document 1/2008 on the protection of Children's Personal Data. http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2008/wp147_en.pdf, feb 2008.

[Art09a] Article 29 Data Protection Working Party. Opinion 2/2009 on the protection of children's personal data (General Guidelines and the special case of schools). http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2009/wp160_en.pdf, jun 2009.

[Art09b] Article 29 Data Protection Working Party. Opinion 5/2009 on online social networking. http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2009/wp160_en.pdf, jun 2009.

[Bar65] Paul Baran. Communications, computers and people. In *Proceedings of Joint Computer Conference, Part II: Computers – their Impact on Society*, AFIPS '65 (Fall, part II), pages 45–49, New York, NY, USA, Dec. 1965. ACM.

[BBP11] Manuela Berg and Katrin Borcea-Pfitzmann. Implementability of the Identity Management Part in Pfitzmann/Hansen's Terminology for a Complex Digital World. In Simone Fischer-Hübner, Marit Hansen, Penny Duquenoy, and Ronald Leenes, editors, *Proceedings of PrimeLife / IFIP Summerschool on Privacy and Identity Management for Life*, IFIP Advances in Information and Communication Technology. Springer, 2011. To be published.

[BCC88] Gilles Brassard, David Chaum, and Claude Crépeau. Minimum disclosure proofs of knowledge. *Journal of Computer and System Sciences*, 37(2):156 – 189, 1988.

[BCC+08] Steve Babbage, Dario Catalano, Carlos Cid, Orr Dunkelman, Christian Gehrmann, Louis Granboulan, Tanja Lange, Arjen Lenstra, Phong Nguyen, Christof Paar, Jan Pelzl, Thomas Pornin, Bart Preneel, Christian Rechberger, Vincent Rijmen, Matt Robshaw, Andy Rupp, Nigel Smart, and Michael Ward. ECRYPT Yearly Report on Algorithms and Keysizes (2007-2008). ECRYPT Deliverable D.SPA.28, ECRYPT Network of Excellence, Aug. 2008.

[BL90] Josh Benaloh and Jerry Leichter. Generalized Secret Sharing and Monotone Functions. In Shafi Goldwasser, editor, *Advances in Cryptology — CRYPTO' 88*, volume 403 of *Lecture Notes in Computer Science*, pages 27–35. Springer, Berlin / Heidelberg, 1990.

[Bla79] G. R. Blakley. Safeguarding cryptographic keys. *International Workshop on Managing Requirements Knowledge*, pages 313–317, 1979.

[BMV06] Johannes Buchmann, Alexander May, and Ulrich Vollmer. Perspectives for cryptographic long-term security. *Communications of the ACM*, 49:50–55, September 2006.

[BRS+09] Rainer Böhme, Maren Raguse, Sandra Steinbrecher, Arnold Roosendaal, Ronald Leenes, Hans Buitelaar, Aleksandra Kuczerawy, Karel Wouters, Marit Hansen, Immanuel Scholz, and Andreas Pfitzmann. Definition of: Prototype ideas for selected scenarios. Primelife Project Heartbeat H1.3.4, PrimeLife Project, Aug. 2009.

[BY97] M. Bellare and B. Yee. Forward integrity for secure audit logs. Technical report, University of California at San Diego, Dept. of Computer Science & Engineering, 1997.

[Cha85] David Chaum. Security without Identification: Transaction Systems to Make Big Brother Obsolete. *Communications of the ACM*, 28(10):1030–1044, 1985.

[CHP+09] Sebastian Clauß, Marit Hansen, Andreas Pfitzmann, Maren Raguse, and Sandra Steinbrecher. Tackling the challenge of lifelong privacy. In P. Cunningham and M. Cunningham, editors, *Proceedings of eChallenges*, Oct. 2009.

[CK01] Sebastian Clauß and Marit Köhntopp. Identity Management and Its Support of Multilateral Security. *Computer Networks*, Special Issue on Electronic Business Systems(37):205–219, 2001. Elsevier, North-Holland.

[CLS11] Jan Camenisch, Ronald Leenes, and Dieter Sommer, editors. *Privacy and Identity Management for Europe (PRIME)*. Springer, 2011.

[Cri99] B. Crispo. Delegation of responsibilities. In Bruce Christianson, Bruno Crispo, William Harbison, and Michael Roe, editors, *Security Protocols*, volume 1550 of *Lecture Notes in Computer Science*, pages 624–625. Springer Berlin / Heidelberg, 1999.

[CSS+05] Jan Camenisch, Abhi Shelat, Dieter Sommer, Simone Fischer-Hübner, Marit Hansen, Henry Krasemann, Gerard Lacoste, Ronald Leenes, and Jimmy Tseng. Privacy and Identity Management for Everyone. In *DIM '05: Proceedings of the 2005 Workshop on Digital Identity Management*, pages 20–27, New York, NY, USA, 2005. ACM.

[DB10] Jaromir Dobiáš and Katrin Borcea-Pfitzmann. Towards a Privacy-Enhanced Backup and Synchronisation Demonstrator Respecting Lifetime Aspects. Primelife Project Heartbeat H1.3.6, PrimeLife Project, March 2010.

[Dom06] Egbert J. Dommering. Regulating technology: code is not law. In E.J. Dommering and L.F. Asscher, editors, *Coding Regulation: Essays on the Normative Role of Information Technology*, pages 1–17, The Hague, 2006. T.M.C. Asser Press.

[DoW08] PrimeLife Description of Work. Annex I to the project proposal. (Internal Document), Version 4 as of February 18, 2008.

[ECH10] Convention for the Protection of Human Rights and Fundamental Freedoms as amended by Protocol No. 11 and No. 14. http://conventions.coe.int/treaty/en/Treaties/Html/005.htm, jun 2010.

[Eur95] European Parliament and Council Directive. Directive 95/46/EC of the European Parliament and of the Council: On the protection of individuals with regard to the processing of personal data and on the free movement of such data. Legal Ruling/Regulation, 1995.

[Eur02] European Parliament and Council Directive. Directive 2002/58/EC of the European Parliament and of the Council: concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications). Official Journal of the European Communities, 2002.

[Eur07] European Commission. Communication from the Commission to the European Parliament and the Council on Promoting Data Protection by Privacy Enhancing Technologies (PETs). http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2007:0228:FIN:EN:PDF, May 2007.

[GMR85] S Goldwasser, S Micali, and C Rackoff. The knowledge complexity of interactive proof-systems. In *Proceedings of the Seventeenth Annual ACM Symposium on Theory of Computing*, STOC '85, pages 291–304, New York, NY, USA, 1985. ACM.

[Gof59] Erving Goffman. *The Presentation of Self in Everyday Life*. Anchor Books, Garden City, New York, June 1959.

[GWB97] I. Goldberg, D. Wagner, and E. Brewer. Privacy-Enhancing Technologies for the Internet. In *Proceedings of the 42nd IEEE International Computer Conference*, pages 103–109. IEEE Computer Society, February 1997.

[GWK+10] Cornelia Graf, Peter Wolkerstorfer, Benjamin Kellermann, Erik Wästlund, and Simone Fischer-Hübner. High-level Prototypes. Primelife Project Deliverable D4.1.4, PrimeLife Project, Aug. 2010.

[HHB+10] Marit Hansen, Leif-Erik Holtz, Hans Buitelaar, Arnold Roosendaal, Aleksandra Kuczerawy, and Karel Wouters. Second thoughts on the WP 1.3 demonstrator. Primelife Project Heartbeat H1.3.7, PrimeLife Project, Oct. 2010.

[HM07] Marit Hansen and Sebastian Meissner, editors. *Verkettung digitaler Identitäten*. Report commissioned by the German Ministry of Education and Research. Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein, Kiel, Dresden, 2007.

[HPS08] Marit Hansen, Andreas Pfitzmann, and Sandra Steinbrecher. Identity Management Throughout One's Whole Life. *Information Security Technical Report*, 13(2):83–94, 2008.

[HRSZ10] Marit Hansen, Maren Raguse, Katalin Storf, and Harald Zwingelberg. Delegation for Privacy Management from Womb to Tomb — A European Perspective. In Michele Bezzi, Penny Duquenoy, Simone Fischer-Hübner, Marit Hansen, and Ge Zhang, editors, *Privacy and Identity Management for Life*, volume 320 of *IFIP Advances in Information and Communication Technology*, pages 18–33. Springer Boston, 2010.

[IDA07] IDABC. eID Interoperability for PEGS: Analysis and Assessment of similarities and differences – Impact on eID interoperability. http://ec.europa.eu/idabc/en/document/6484/5938, Nov. 2007.

[KDR+08] Eleni Kosta, Jos Dumortier, Piet Ribbers, Alea Fairchild, Jimmy Tseng, Katja Liesbach, Elke Franz, Ronald Leenes, Marcel Hoogwout, Bart Priem, Tobias Kölsch, Jan Zibuschka, Georg Kramer, and Günter Schumacher. Requirements for Privacy Enhancing Tools. PRIME Project Deliverable D1.1.d, Prime Project, March 2008.

[Kir08] Michael Kirby. Law Making Meets Technology. http://www.onlineopinion.com.au/view.asp?article=7082&page=0, 2008.

[Kor09] Douwe Korff. Are Users of Social Networking Sites Subject to Data Protection Law, as Controllers? *dataprotectionreview.eu*, (9):4, June 2009.

[Kun07] Christopher Kuner. *European Data Protection Law: Corporate Regulation and Compliance*. Oxford University Press, USA, 2 edition, 4 2007.

[LSH08] Ronald Leenes, Jan Schallaböck, and Marit Hansen. Prime white paper v3. Technical report, PRIME Project, May 15 2008.

[Mei09] Martin Meints. The Relationship between Data Protection Legislation and Information Security Related Standards. In Vashek Matyáš, Simone Fischer-Hübner, Daniel Cvrcek, and Petr Švenda, editors, *The Future of Identity in the Information Society*, volume 298 of *IFIP Advances in Information and Communication Technology*, pages 254–267. Springer Boston, 2009.

[MS07] Viktor Mayer-Schönberger. Useful Void: The Art of Forgetting in the Age of Ubiquitous Computing. Working Paper Series RWP07-022, Harvard University, John F. Kennedy School of Government, Apr. 2007.

[OEC80] Recommendation of the Council Concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data. http://www.oecd.org/document/18/0,3343,en_2649_34255_1815186_1_1_1_1,00.html, Sept. 23, 1980.

[PBP10] Andreas Pfitzmann and Katrin Borcea-Pfitzmann. Lifelong Privacy: Privacy and Identity Management for Life. In Michele Bezzi, Penny Duquenoy, Simone Fischer-Hübner, Marit Hansen, and Ge Zhang, editors, *Privacy and Identity Management for Life*, volume 320/2010 of *IFIP Advances in Information and Communication Technology*, pages 1–17, Boston, 2010. Springer.

[PH10] Andreas Pfitzmann and Marit Hansen. A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management. https://dud.inf.tu-dresden.de/Anon_Terminology.shtml, August 2010. v0.34.

[Phi04] David J. Phillips. Privacy policy and PETS – The influence of policy regimes on the development and social implications of privacy enhancing technologies. *New Media & Society*, 6(6):691 – 706, 2004.

[Pov00] Dean Povey. Optimistic Security: A New Access Control Paradigm. In *Proceedings of the 1999 workshop on New security paradigms*, NSPW '99, pages 40–45, New York, NY, USA, 2000. ACM.

[PRMD10] Quan Pham, Jason Reid, Adrian McCullagh, and Edward Dawson. On a Taxonomy of Delegation. *Computers & Security*, 29(5):565 – 579, 2010. Challenges for Security, Privacy and Trust.

[PSDCP08] R. Peeters, K. Simoens, D. De Cock, and B. Preneel. Cross-Context Delegation through Identity Federation. In Arslan Brömme, Christoph Busch, and Detlef Hühnlein, editors, *Proceedings of BIOSIG: Biometrics and Electronic Signatures*, number P-137 in Lecture Notes in Informatics, pages 79–92. Bonner Köllen Verlag, 2008.

[Rac75] James Rachels. Why Privacy is Important. *Philosophy & Public Affairs*, 4(4):323–333, 1975.

[RBB+08] M. Roussopoulos, L. Beslay, C. Bowden, G. Finocchiaro, M. Hansen, M. Langheinrich, G. Le Grand, and K. Tsakona. *Technology-Induced Challenges in Privacy and Data Protection in Europe*. enisa – European Network and Information Security Agency, Oct. 2008.

[RSA78] R. L. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 21:120–126, Feb. 1978.

[RSH+09] Arnold Roosendaal, Sandra Steinbrecher, Hans Hedbom, Stuart Short, Aleksandra Kuczerawy, and Maren Raguse. Analysis of privacy and identity management throughout life. Primelife Project Heartbeat H1.3.3, PrimeLife Project, June 2009.

[SA08] William Seltzer and Margo Anderson. Using population data systems to target vulnerable population subgroups and individuals: Issues and incidents. In Jana Asher, David Banks, and Fritz J. Scheuren, editors, *Statistical Methods for Human Rights*, pages 273–328. Springer New York, 2008.

[Sca96] Teresa Scassa. National Identity, Ethnic Surnames and the State. *Canadian Journal of Law and Society*, 11(2):167–191, 1996.

[Sha79] Adi Shamir. How to share a secret. *Communications of the ACM*, 22:612–613, Nov. 1979.

[SHP+09] Katalin Storf, Marit Hansen, Andreas Pfitzmann, Sandra Steinbrecher, Arnold Roosendaal, Ulrich Pinsdorf, Karel Wouters, Aleksandra Kuczerawy, Rainer Böhme, and Stefan Berthold. Requirements and concepts for identity management throughout life. Primelife Project Heartbeat H1.3.5, PrimeLife Project, Nov. 2009.

[Sim06] Spiros Simitis. *Bundesdatenschutzgesetz*. Nomos, Frankfurt, 2006.

[SK99] Bruce Schneier and John Kelsey. Secure audit logs to support computer forensics. *ACM Transactions on Information and System Security (TISSEC)*, 2(2):159–176, May 1999.

[SRS+98] C. Soutar, D. Roberge, A. Stoianov, R. Gilroy, and B. Kumar. Biometric encryption. In Randall K. Nichols, editor, *ICSA Guide to Cryptography*, pages 649–6750. McGraw-Hill, 1998.

[SW05] Amit Sahai and Brent Waters. Fuzzy identity-based encryption. In Ronald Cramer, editor, *Advances in Cryptology – EUROCRYPT 2005*, volume 3494 of *Lecture Notes in Computer Science*, pages 457–473. Springer, Berlin / Heidelberg, 2005.

[vdBL10] Bibi van den Berg and Ronald Leenes. Audience Segregation in Social Network Sites. In *IEEE Second International Conference on Social Computing (SocialCom)*, pages 1111 –1116, 2010.

[Whi04] James Q. Whitman. The Two Western Cultures of Privacy: Dignity versus Liberty. *Yale Law Journal*, 113, 2004.

[ZFK+98] J. Zöllner, H. Federrath, H. Klimant, A. Pfitzmann, R. Piotraschke, A. Westfeld, G. Wicke, and G. Wolf. Modeling the Security of Steganographic Systems. In *Proceedings of Second International Workshop on Information Hiding 1998*, volume 1525 of *Lecture Notes in Computer Science*, pages 344–354. Berlin / Heidelberg, 1998.